Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies

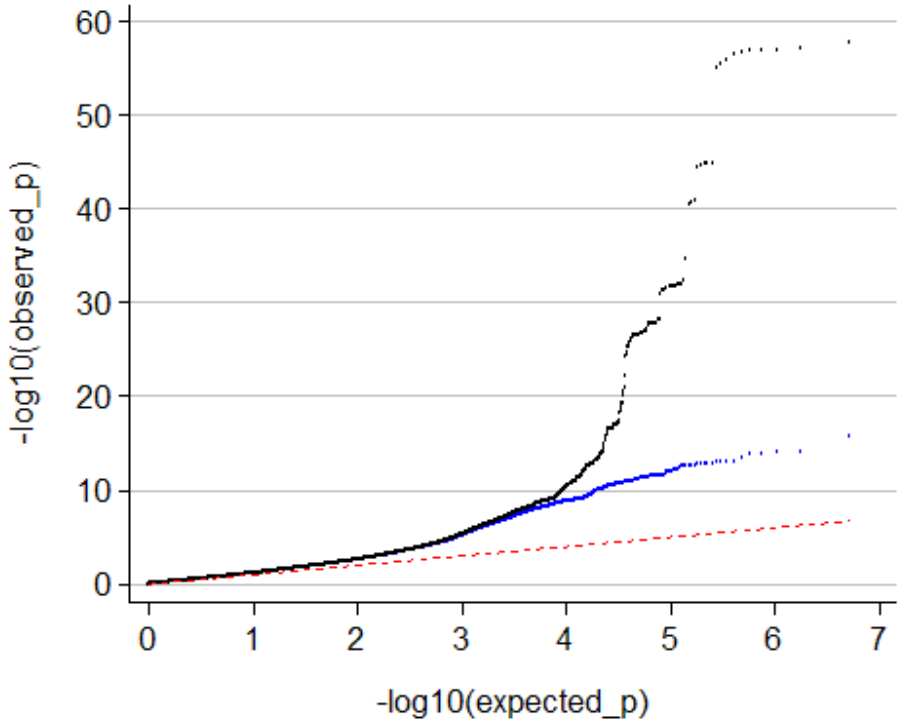
Cathy E. Elks*[1], John R.B. Perry*[2], Patrick Sulem*[3], Daniel I. Chasman[4,5], Nora Franceschini[6], Chunyan He[7,8], Kathryn L. Lunetta[9,10], Jenny A.Visser[11], Enda M. Byrne[12,13], Diana L. Cousminer[14], Daniel F. Gudbjartsson[3], Tõnu Esko[15,16,17], Bjarke Feenstra[18], Jouke-Jan Hottenga[19], Daniel L. Koller[20], Zoltán Kutalik[21,22], Peng Lin[23], Massimo Mangino[24], Mara Marongiu[25], Patrick F. McArdle[26], Albert V. Smith[27,28], Lisette Stolk[11,29], Sophie W. van Wingerden[30], Jing Hua Zhao[1], Eva Albrecht[31], Tanguy Corre[32], Erik Ingelsson[33], Caroline Hayward[34], Patrik K.E. Magnusson[33], Erin N. Smith[35], Shelia Ulivi[36], Nicole M. Warrington[37], Lina Zgaga[38], Helen Alavere[15], Najaf Amin[30], Thor Aspelund[27,28], Stefania Bandinelli[39], Ines Barroso[40], Gerald S. Berenson[41], Sven Bergmann[21,22], Hannah Blackburn[40], Eric Boerwinkle[42], Julie E. Buring[4,43,] Fabio Busonero[25], Harry Campbell[38], Stephen J. Chanock[44], Wei Chen[41], Marilyn C. Cornelis[45], David Couper[46], Andrea D. Coviello[47], Pio d'Adamo[36], Ulf de Faire[48], Eco J.C. de Geus[19], Panos Deloukas[40], Angela Döring[31], George Davey Smith[49], Douglas F. Easton[50], Gudny Eiriksdottir[27], Valur Emilsson[51], Johan Eriksson[52,53,54,55], Luigi Ferrucci[56], Aaron R. Folsom[57], Tatiana Foroud[20], Melissa Garcia[58], Paolo Gasparini[36], Frank Geller[18], Christian Gieger[31], The GIANT Consortium[59], Vilmundur Gudnason[27,28], Per Hall[33], Susan E. Hankinson[43,60], Liana Ferreli[25], Andrew C. Heath[61], Dena G. Hernandez[62], Albert Hofman[63], Frank B. Hu[43,45,60], Thomas Illig[31], Marjo-Riitta Järvelin[64], Andrew D. Johnson[9,65], David Karasik[66], Kay-Tee Khaw [67], Douglas P. Kiel[66], Tuomas O. Kilpeläinen[1], Ivana Kolcic[68], Peter Kraft[43,45,60], Lenore J. Launer[58], Joop S.E. Laven[69], Shengxu Li[1], Jianjun Liu[70], Daniel Levy[9, 65 ,71], Nicholas G. Martin[72], Wendy L. McArdle[73], Mads Melbye[18], Vincent Mooser[74], Jeffrey C. Murray[75], Sarah S. Murray[35], Michael A. Nalls[76], Pau Navarro[34], Mari Nelis[15,16,17], Andrew R. Ness[77], Kate Northstone[73], Ben A. Oostra[30], Munro Peacock[78], Lyle J. Palmer[37], Aarno Palotie[14,40,79], Guillaume Paré[4,5 ,80], Alex N. Parker[81], Nancy L. Pedersen[33], Leena Peltonen[14,40,52,79, 82,] Craig E. Pennell[83], Paul Pharoah[50], Ozren Polasek[68,84], Andrew S. Plump[85], Anneli Pouta[52], Eleonora Porcu[25], Thorunn Rafnar[3], John P. Rice[23], Susan M. Ring[73], Fernando Rivadeneira[11,29,63], Igor Rudan[38,86], Cinzia Sala[32], Veikko Salomaa[52], Serena Sanna[25], David Schlessinger[87], Nicholas J. Schork[35], Angelo Scuteri[25,88], Ayellet V. Segrè[79,89], Alan R. Shuldiner[26,90], Nicole Soranzo[24,40], Ulla Sovio[64], Sathanur R. Srinivasan[41], David P. Strachan[91], Mar-Liis Tammesoo[15], Emmi Tikkanen[14,.52], Daniela Toniolo[32], Kim Tsui[81], Laufey Tryggvadottir[92], Jonathon Tyrer[50], Manuela Uda[25], Rob M. van Dam[45,93], Joyve B.J. van Meurs[11], Peter Vollenweider[94], Gerard Waeber[94], Nicholas J. Wareham[1], Dawn M. Waterworth[74], Michael N. Weedon[2], H. Erich Wichmann[31,95,96], Gonneke Willemsen[19], James F. Wilson[38], Alan F. Wright[34], Lauren Young[81], Guangju Zhai[24], Wei Vivian Zhuang[10], Laura J. Bierut[23], Dorret I. Boomsma[19], Heather A. Boyd[18], Laura Crisponi[25], Ellen W. Demerath[57], Cornelia M. van Duijn[30], Michael J. Econs[20, 78], Tamara B. Harris[58], David J. Hunter[43,44,45,60], Ruth J.F. Loos[1], Andres Metspalu[15,16,17], Grant W. Montgomery[97], Paul M. Ridker[4,5,43,98,] Tim D. Spector[24], Elizabeth A. Streeten[26], Kari Stefansson[3,99], Unnur Thorsteinsdottir[3,99], André G. Uitterlinden[11,29,63], Elisabeth Widen[14], Joanne M. Murabito**[9,47], Ken K. Ong**[1,100], Anna Murray**[2]

SUPPLEMENTARY INFORMATION

SUPPLEMENTARY FIGURE LEGENDS

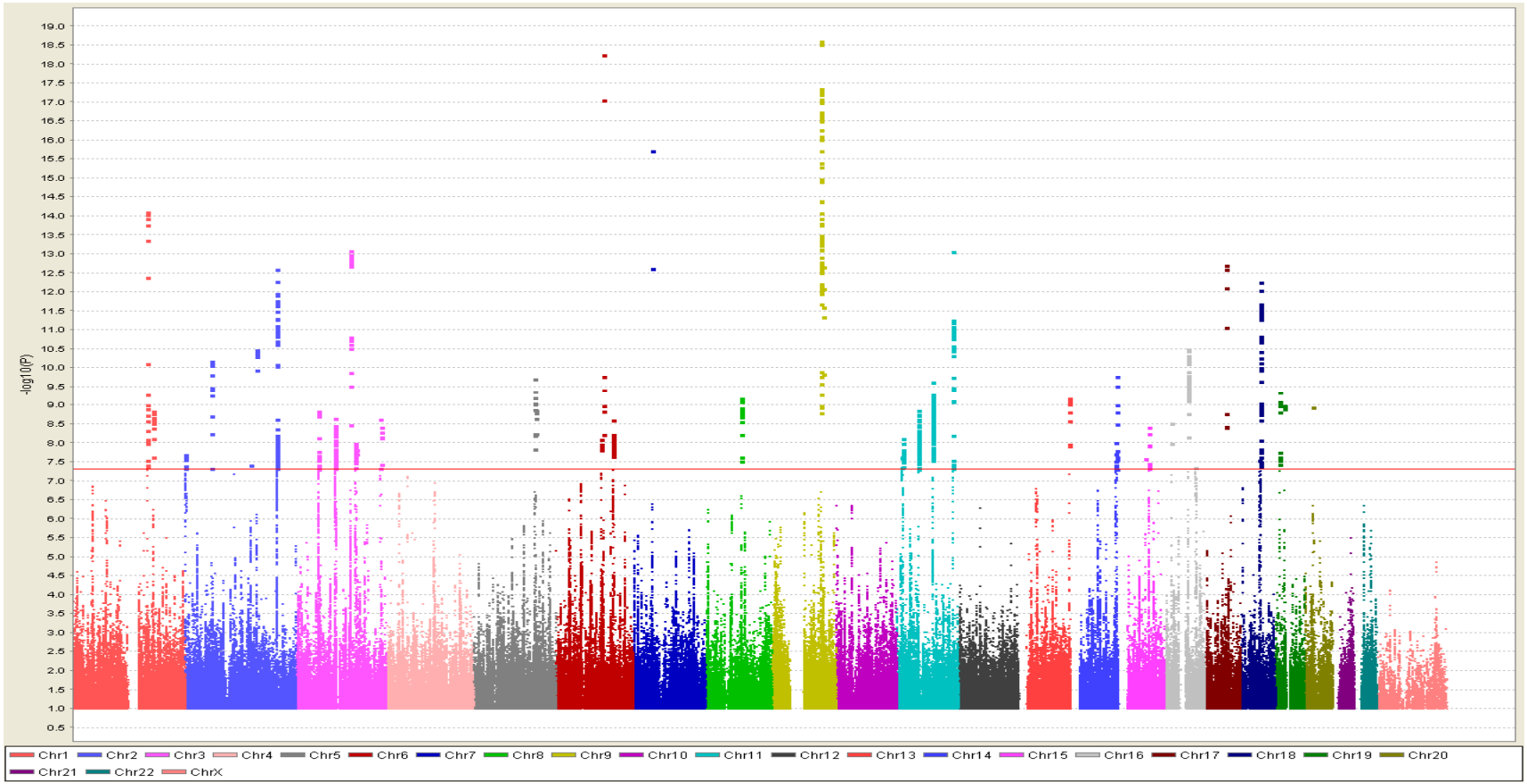
Supplementary Fig. 1: Quantile-quantile plots of the genome-wide scan for age at menarche.

The dots represent the observed –log10 *P*-values before (black dots) and after (blue dots) removal of signals associated with the two previously identified menarche loci at *LIN28B* and 9q31.2. The expected distribution of –log10 *P* -values under the null hypothesis is shown by the dashed red line.



Supplementary Fig. 2: Manhattan plot of GWAS for age at menarche from the Phase 1 meta-analysis of 32 studies.

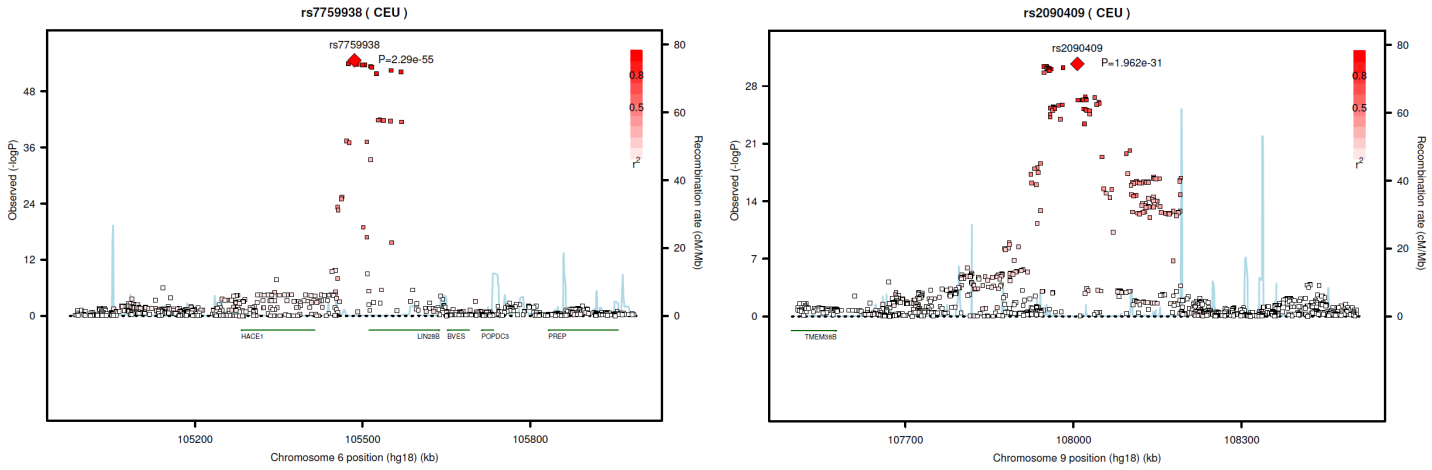
X-axis represents the chromosomal position for each SNP, and y-axis the –log10 P-value for association with age at menarche. Note: To highlight the novel menarche loci the y-axis is curtailed at –log10 P-value = 20, and therefore the strongest signals at the two previously identified loci on Chromosomes 6 and 9 are not displayed.



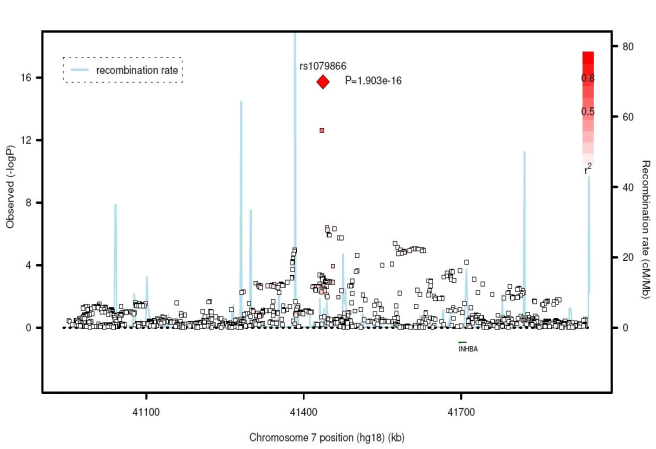
Supplementary Fig 3: Regional association plots for each of the 42 known, confirmed or possible novel menarche loci.

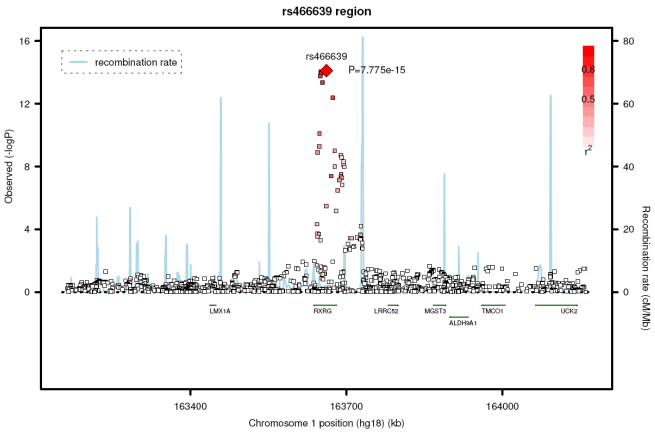
SNPs are plotted by chromosomal position (x-axis) against GWAS association with age at menarche (–log10 P value). The strongest signal is denoted by the figure sub-title and red diamond. Other SNPs are colour coded to reflect their LD with the top SNP (taken from pairwise r2 values from the HapMap CEU database). Estimated recombination rates (from HapMap) are plotted in cyan to reflect the local LD structure. Genes are denoted by green lines. Figures were drawn using SNAP LD (http://www.broadinstitute.org/mpg/snap/index.php#citation) except where indicated.

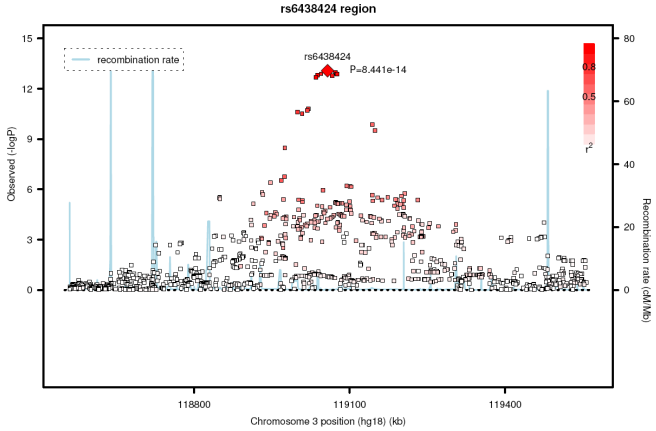
Two known menarche loci:

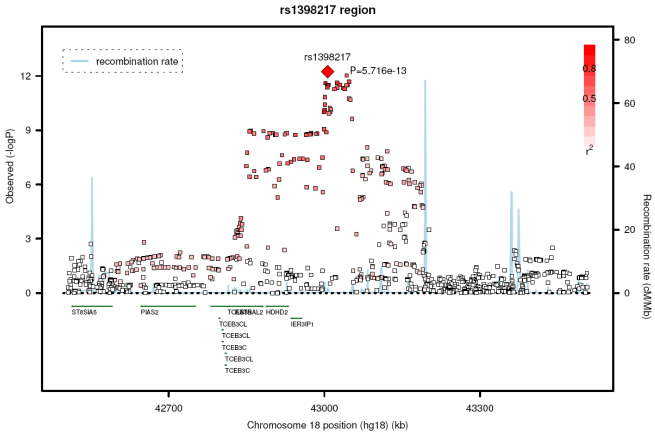


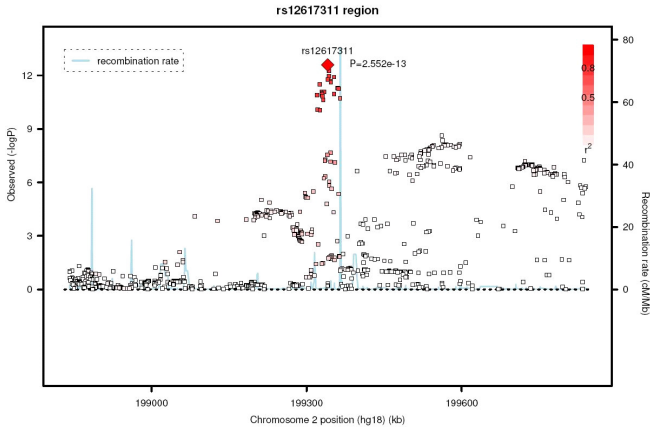
Thirty confirmed novel menarche loci:

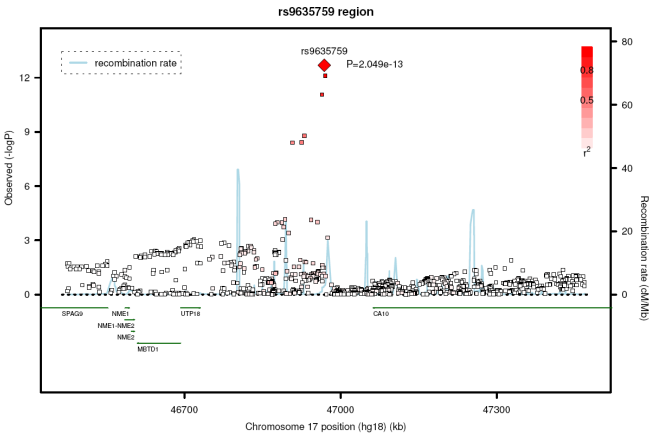


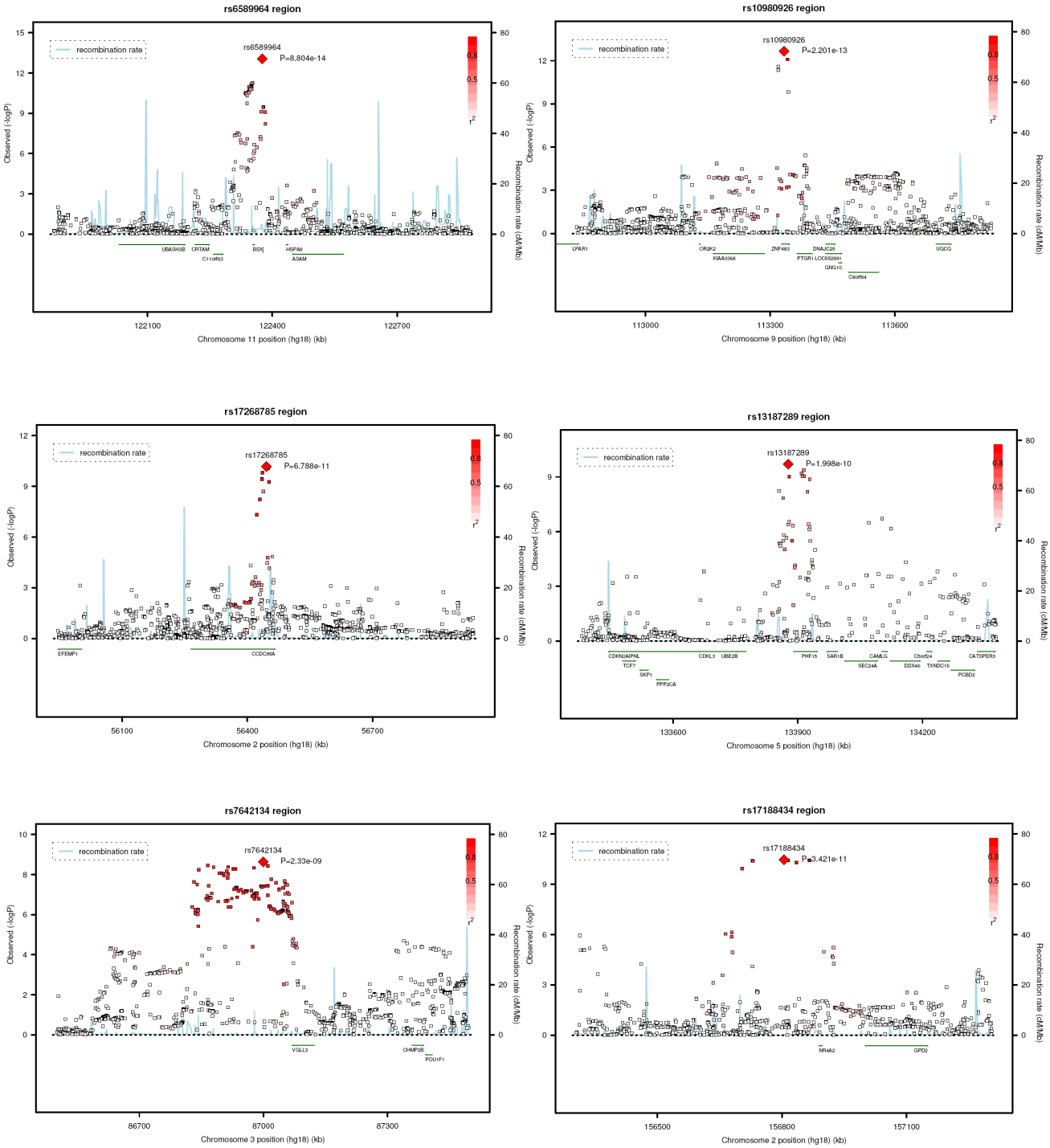


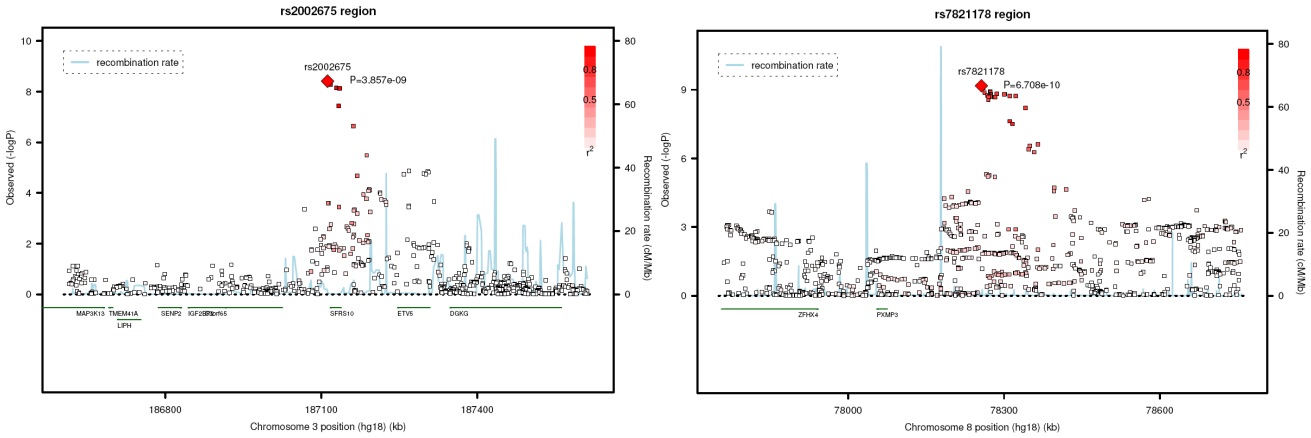


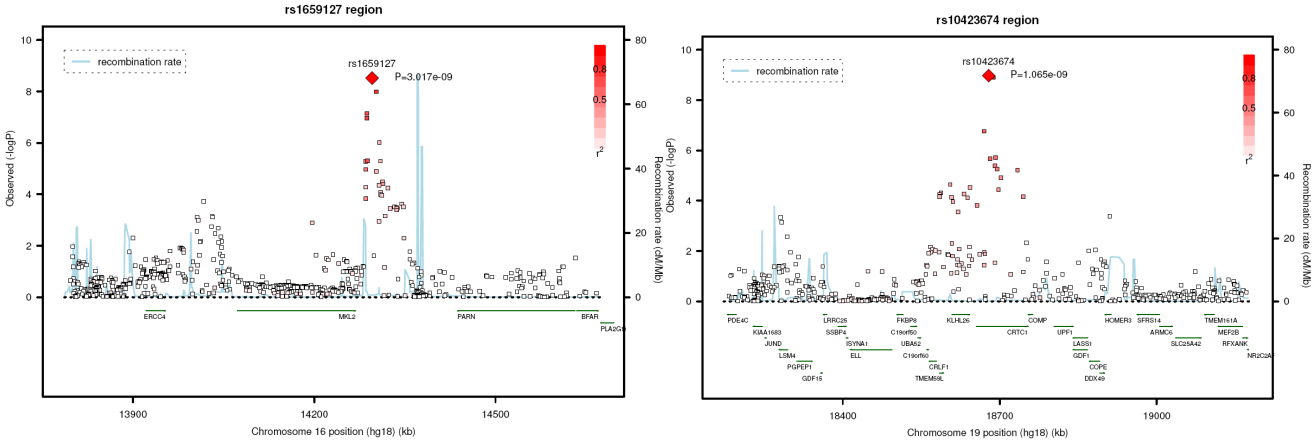


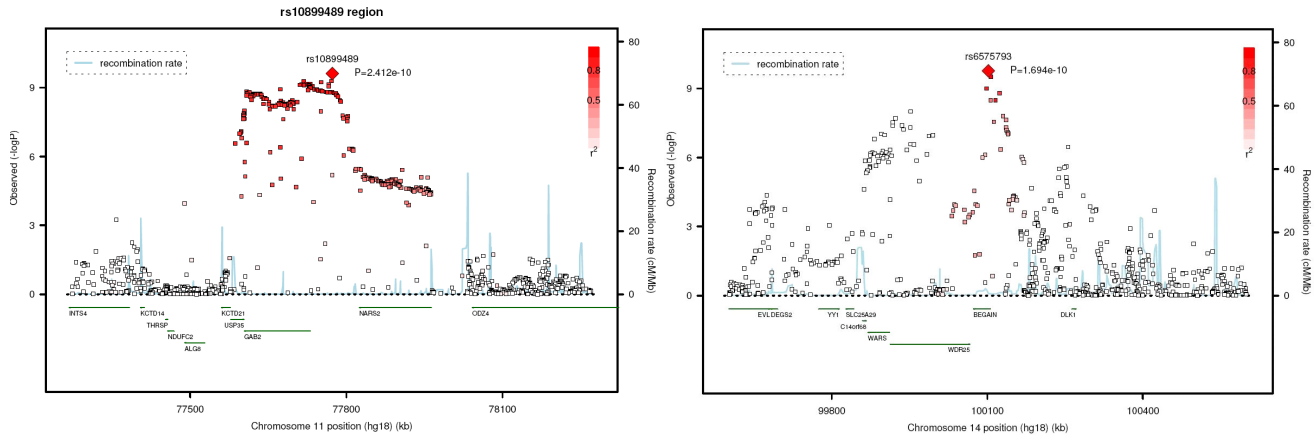




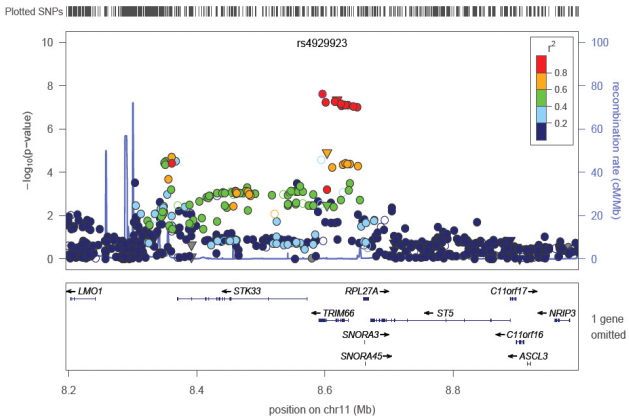

rs6589964 region
recombination rate
rs6589964
P=8.804e-14
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 11 position (hg18) (kb)
rs10980926 region
rs10980926
P=2.201e-13
Chromosome 9 position (hg18) (kb)
rs17268785 region
rs17268785
P=6.788e-11
Chromosome 2 position (hg18) (kb)
rs13187289 region
rs13187289
P=1.998e-10
Chromosome 5 position (hg18) (kb)
rs7642134 region
rs7642134
P=2.33e-09
Chromosome 3 position (hg18) (kb)
rs17188434 region
rs17188434
P=3.421e-11
Chromosome 2 position (hg18) (kb)

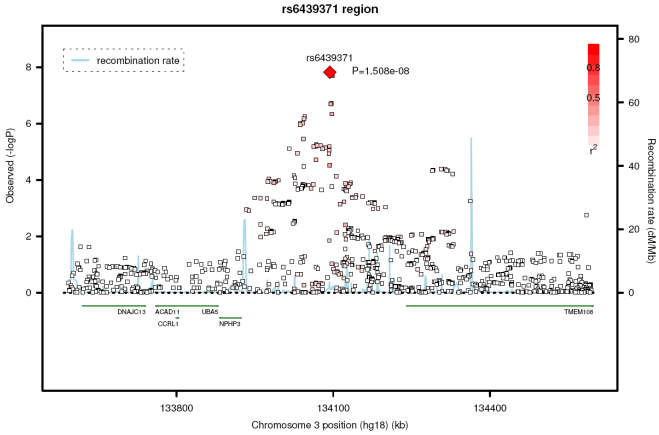

rs2002675 region
recombination rate
rs2002675
P=3.857e-09
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 3 position (hg18) (kb)
rs7821178 region
rs7821178
P=6.708e-10
Chromosome 8 position (hg18) (kb)

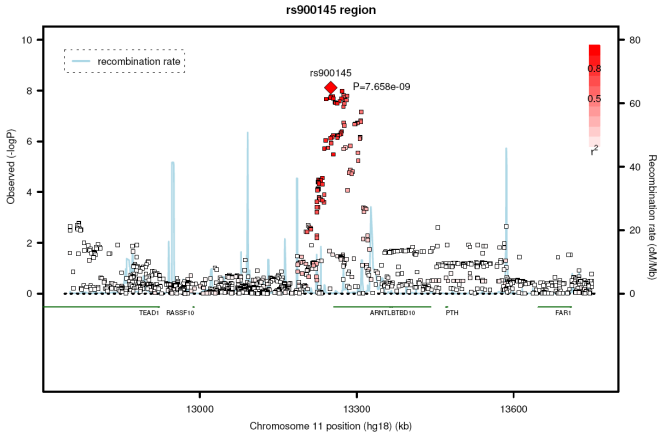

rs1659127 region
rs1659127
P=3.017e-09
Chromosome 16 position (hg18) (kb)
rs10423674 region
rs10423674
P=1.065e-09
Chromosome 19 position (hg18) (kb)

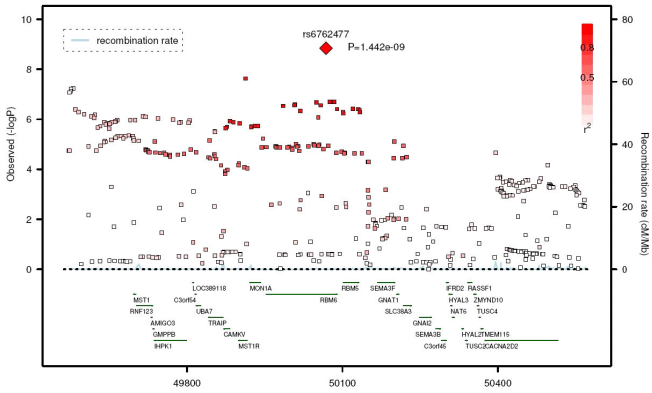

rs10899489 region
rs10899489
P=2.412e-10
Chromosome 11 position (hg18) (kb)
rs6575793
P=1.694e-10
Chromosome 14 position (hg18) (kb)

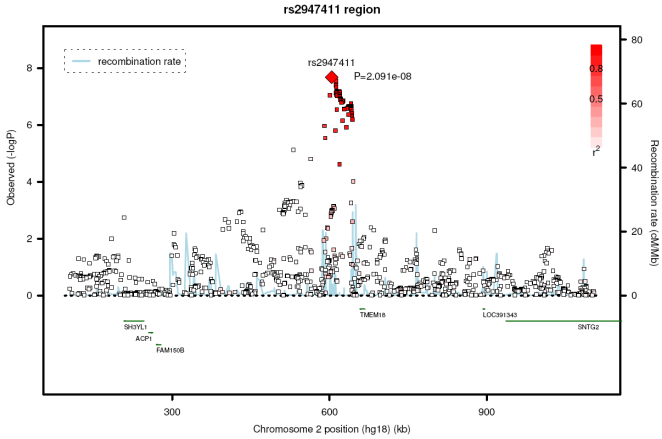
The rs4929923 region would not upload to SNAP and was drawn using LocusZoom (http://csg.sph.umich.edu/locuszoom/)

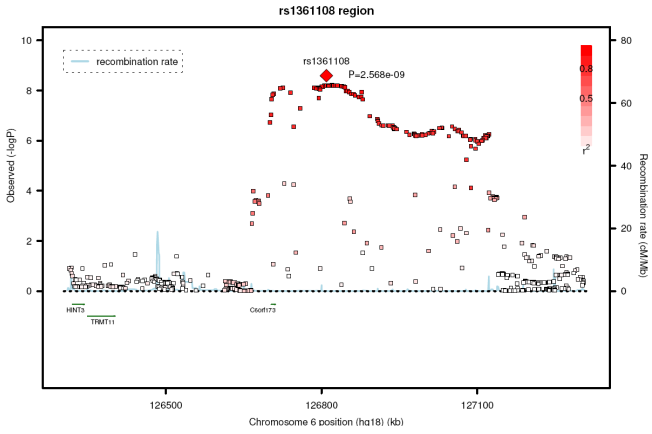


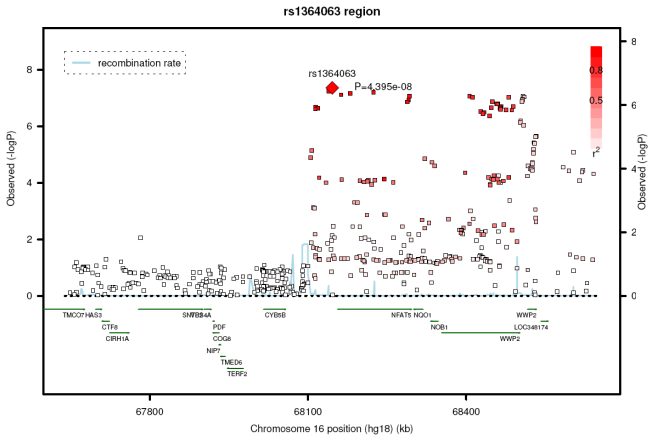


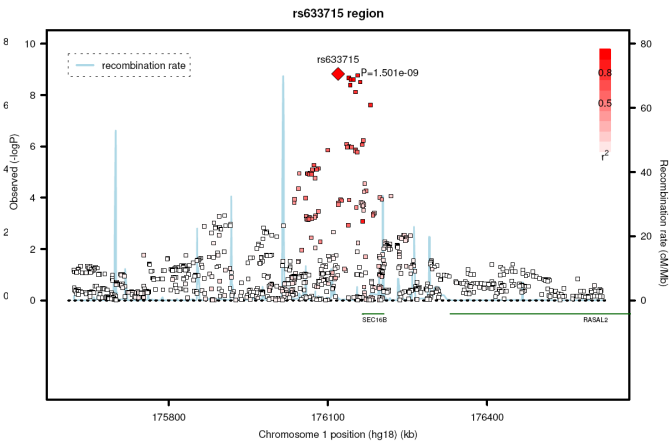


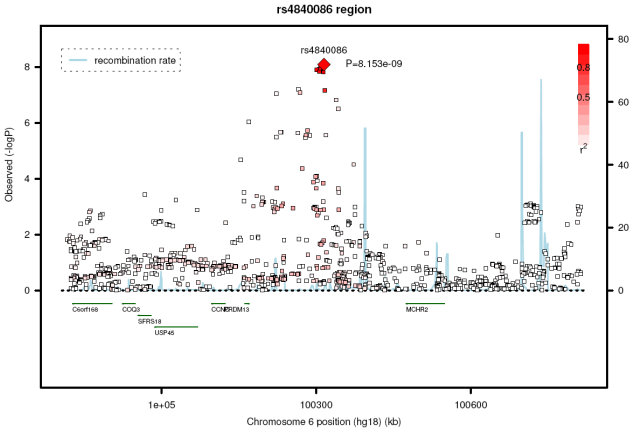


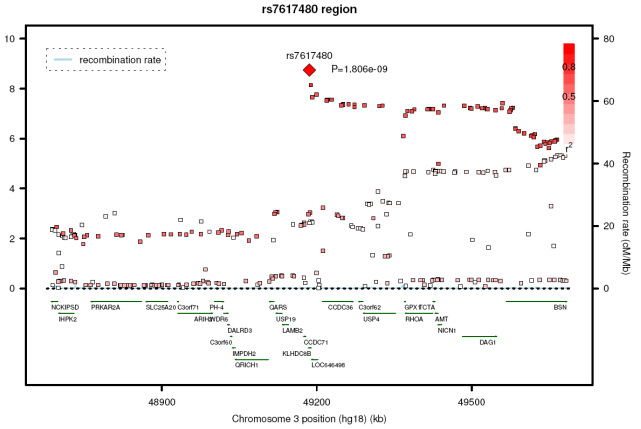


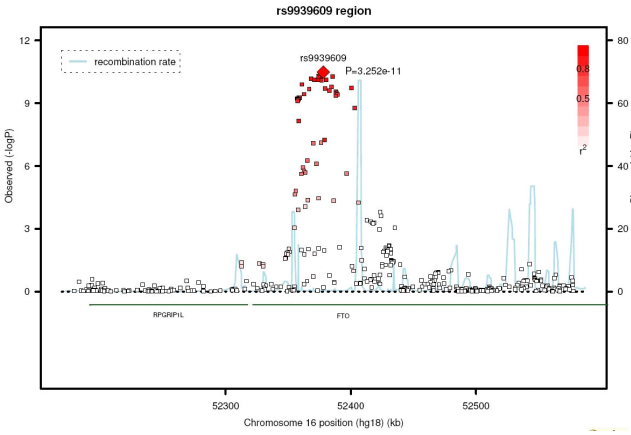


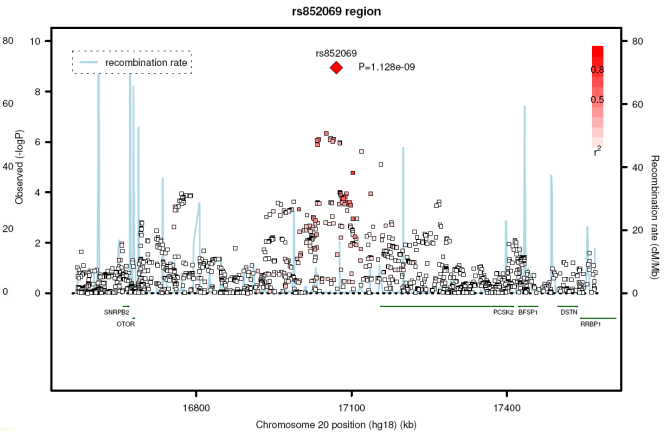

rs1364063 region
recombination rate
rs1364063
P=4.395e-08
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 16 position (hg18) (kb)


rs633715 region
recombination rate
rs633715
P=1.501e-09
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 1 position (hg18) (kb)

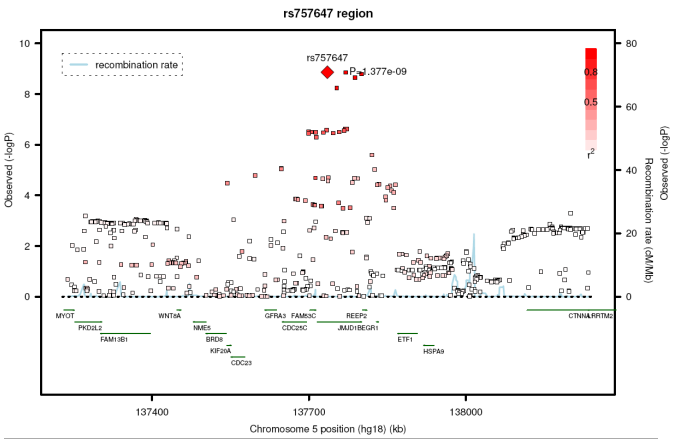

rs4840086 region
recombination rate
rs4840086
P=8.153e-09
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 6 position (hg18) (kb)

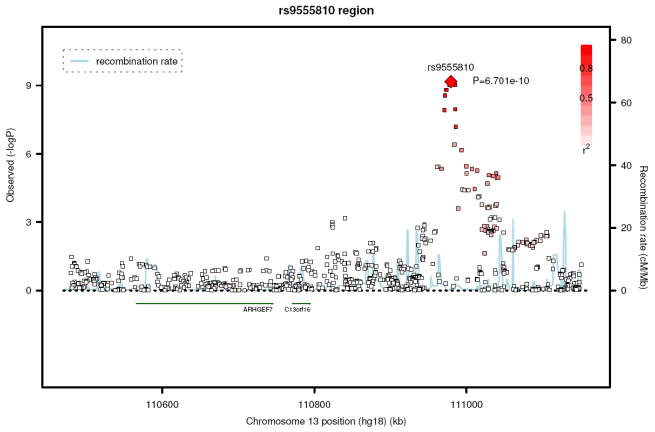

rs7617480 region
recombination rate
rs7617480
P=1.806e-09
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 3 position (hg18) (kb)

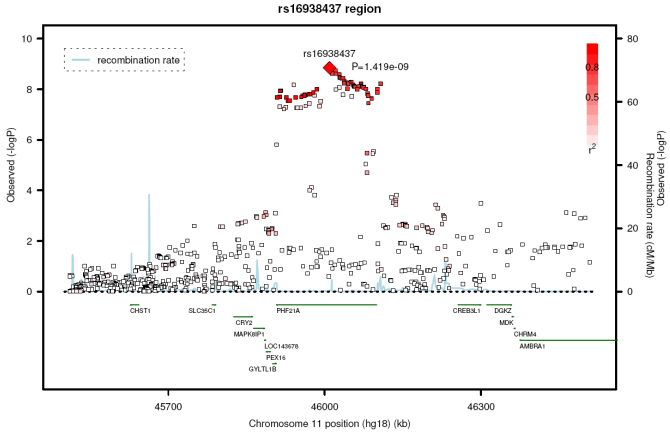

rs9939609 region
recombination rate
rs9939609
P=3.252e-11
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 16 position (hg18) (kb)

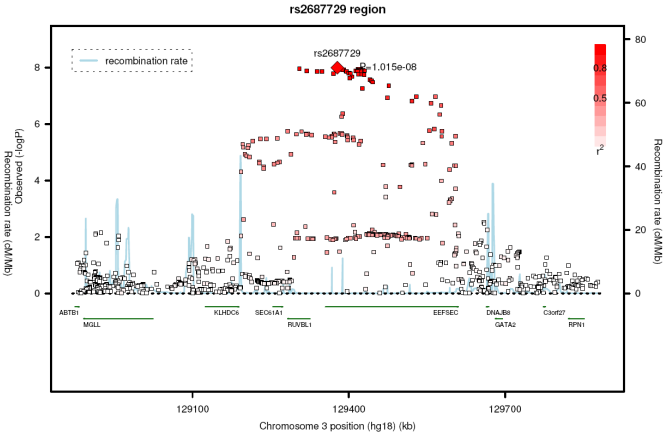

rs852069 region
recombination rate
rs852069
P=1.128e-09
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 20 position (hg18) (kb)

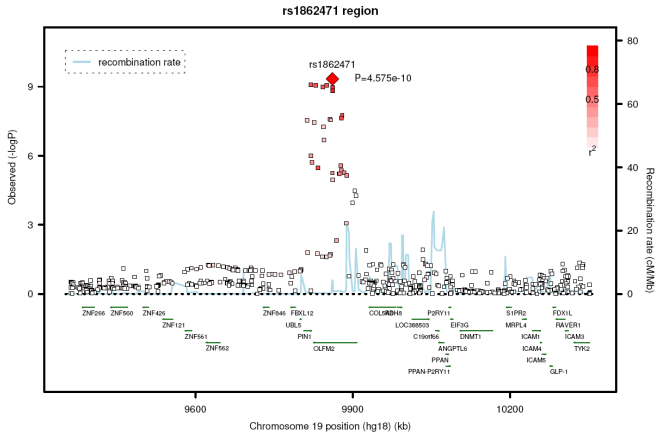
Ten possible novel menarche loci:

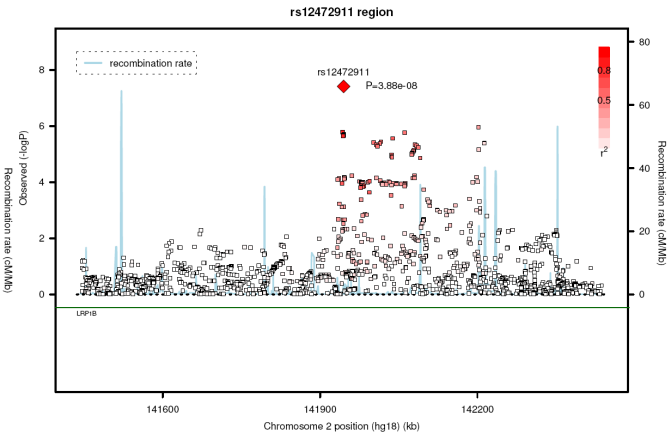


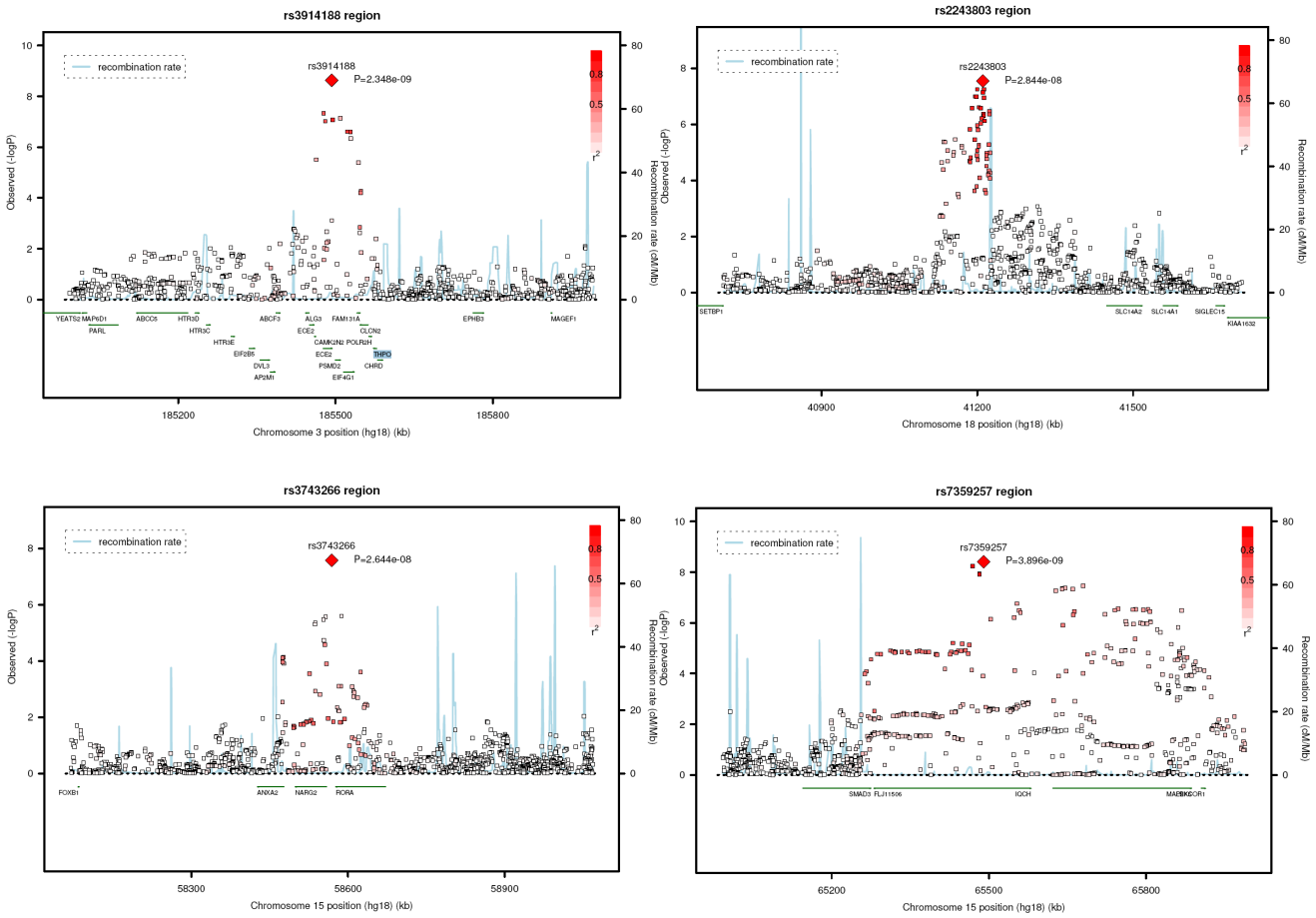






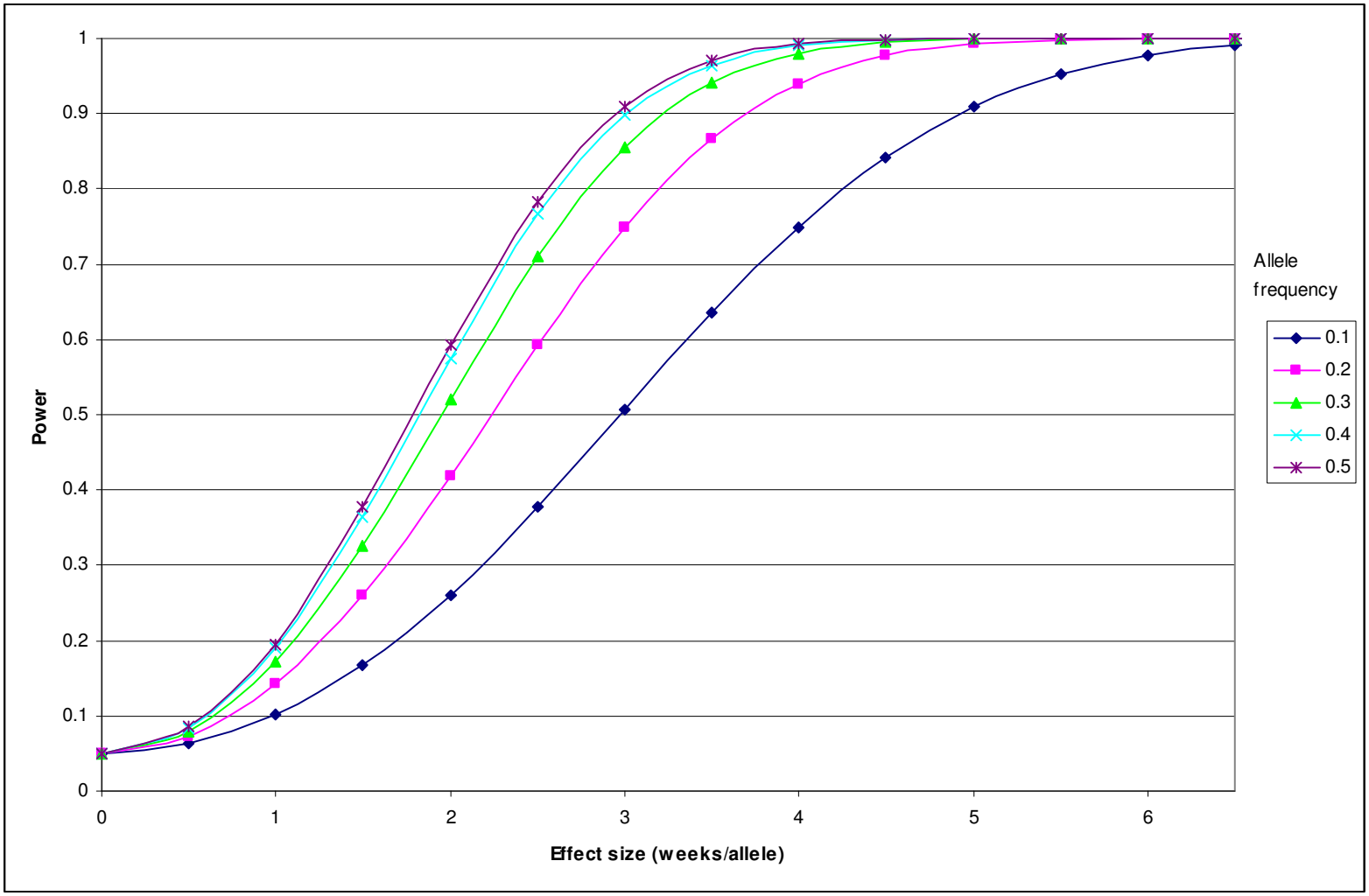





rs3914188 region
recombination rate
rs3914188
P=2.348e-09
Observed (-logP)
Recombination rate (cM/Mb)
Chromosome 3 position (hg18) (kb)
rs2243803 region
rs2243803
P=2.844e-08
Chromosome 18 position (hg18) (kb)
rs3743266 region
rs3743266
P=2.644e-08
Chromosome 15 position (hg18) (kb)
rs7359257 region
rs7359257
P=3.896e-09
Chromosome 15 position (hg18) (kb)

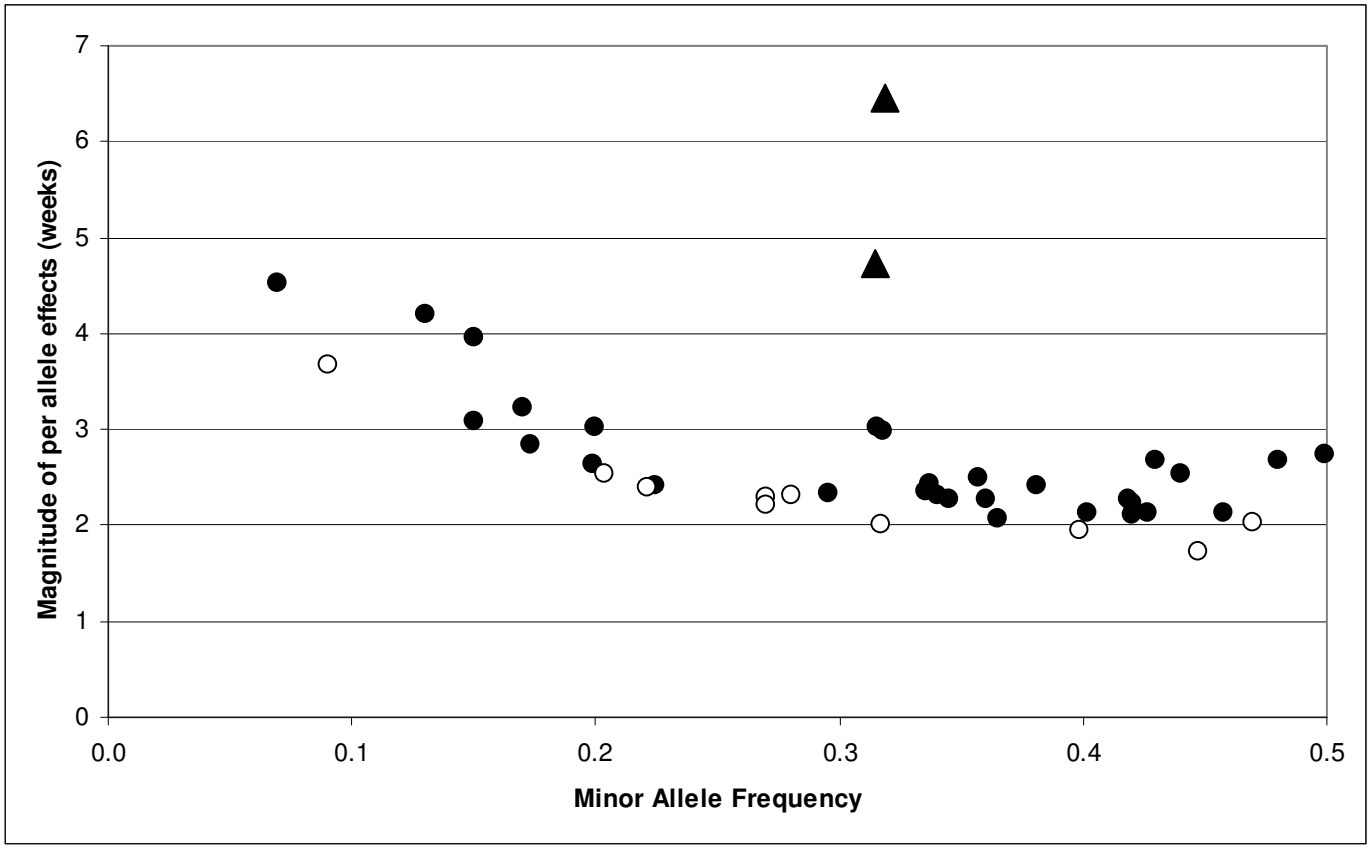
Supplementary Fig. 4: Statistical power of the age at menarche replication studies (n=14,731 women).

Based on age at menarche as a continuous outcome with mean ± SD 13 ± 1.5 years, additive models and 2-sided P<0.05.



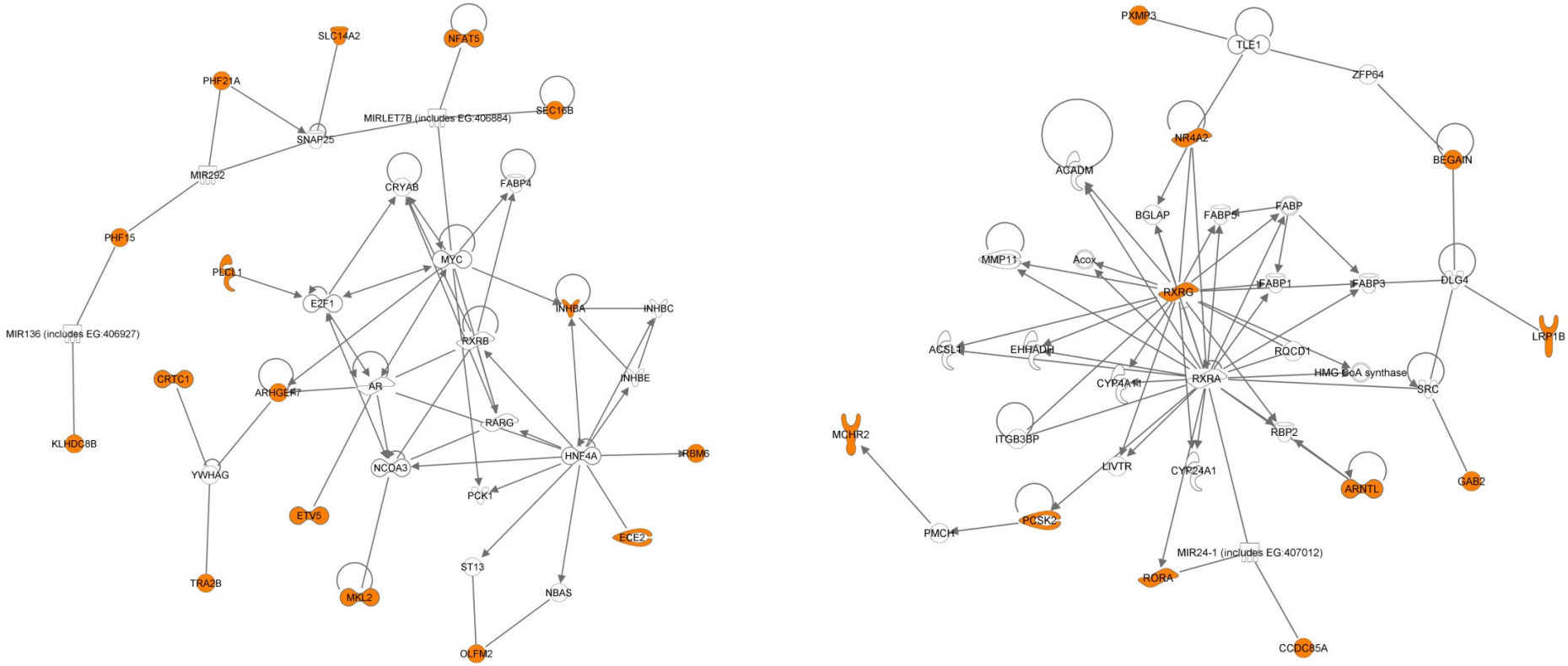
Supplementary Fig. 5: Estimated magnitude of per allele effects of the 42 known, confirmed or possible novel menarche loci plotted by minor allele frequencies.

The two known loci are indicated by triangles, the 30 confirmed novel loci by filled circles and the 10 possible novel loci by open circles.



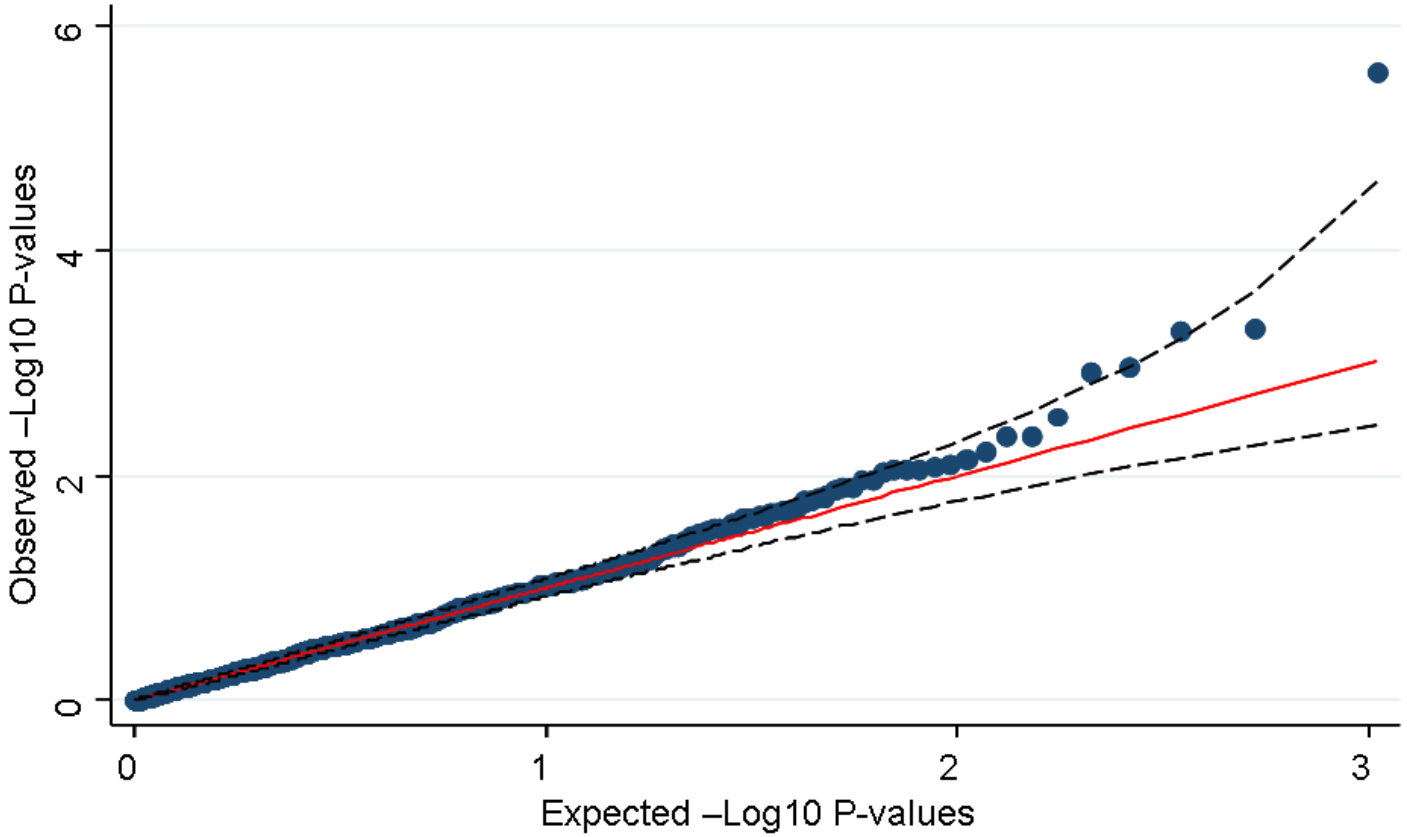
Supplementary Fig. 6A and 6B: Networks revealed through the Ingenuity Pathway Analysis.

Based on genes nearest to the 42 known, confirmed or possible novel menarche loci. A) Network 1, $p=1x10^{-37}$; B) Network 2, $p=1x10^{23}$. Molecules highlighted in orange represent the genes nearest to the menarche loci. Grey lines indicate the direct relationship between genes and molecules.



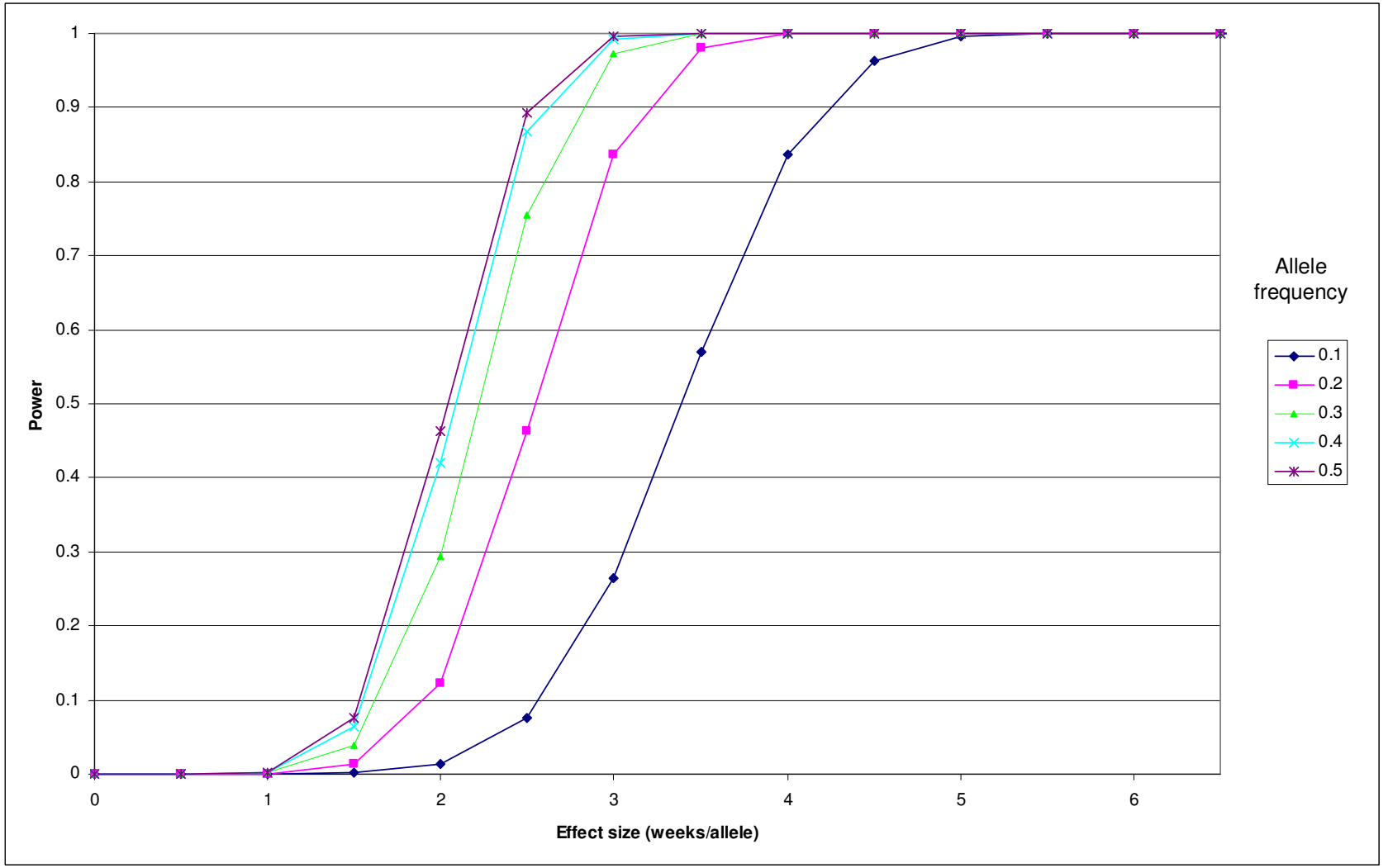
Supplementary Fig. 7: Quantile-quantile plot from the CNV tagging SNP scan.

Based on the list of CNV tagging SNPs (N=1052) from Conrad et al. Nature 2009 (Ref 25)
The only hit which survived multiple testing is at *NEGR1*, the known CNV locus for BMI (rs3101336, MAF 0.39).



Supplementary Fig. 8: Statistical power of the phase 1 meta analysis for age at menarche (n=87,802 women).

Based on age at menarche as a continuous outcome with mean ± SD 13 ± 1.5 years, additive models and 2-sided $P<5x10^{-8}$.



Supplementary Table 1: Details of the 32 studies that contributed to the Stage 1 GWAS meta-analysis with age at menarche.

Study name/acronym	*Full study name*	*N*	GC Factor	*Mean age (SD)*	***Mean AAM (SD)***	*Specific menarche questions*
AGES-Reykjavik	Age, Gene/Environment Susceptibility Study	1849	1.03	76.3 (5.5)	13.6 (1.3)	*"At what age did your menstrual periods begin?"*
Amish	The older order Amish population studies	557	1.046	49.1 (3.7)	13.1 (1.3)	*"How old were you when you had your first menstrual period?"*
ARIC	Atherosclerosis Risk in Communities Study	4247	1.028	53.9 (5.7)	12.9 (1.5)	*"At approximately what age were you when your menstrual periods started?"*
B58C-T1DGC	British 1958 birth cohort (Type 1 Diabetes Genetic Consortium controls)	1021	1.018	16.1 (0.2)	12.7 (1.4)	*Question asked to parents (usually mother): "At what age did she have her first menstrual period?" Answers were coded as "before 11th birthday" (we coded as 10), "when aged 11", "aged 12", "aged 13", "aged 14", "aged 15 or more", "not yet commenced" (which we coded as 16), "commenced but don't know when" (excluded), "don't know if commenced" (excluded).*
B58C-WTCCC	British 1958 birth cohort (Wellcome Trust Case Control Consortium controls)	563	0.998	16.1 (0.2)	12.8 (1.3)	*Question asked to parents (usually mother): "At what age did she have her first menstrual period?" Answers were coded as "before 11th birthday" (we coded as 10), "when aged 11", "aged 12", "aged 13", "aged 14", "aged 15 or more", "not yet commenced" (which we coded as 16), "commenced but don't know when" (excluded), "don't know if commenced" (excluded).*
CoLaus	Cohort Lausannoise	2797	1.03	53.4 (10.8)	13.2 (1.6)	*At what age did you have your first period?*
deCODE	deCODE Genetics, Iceland	15,864	1.284	birth year 1948.1 (17.0)	13.2 (1.3)	*How old were you when your menstruation started?*
DNBC	Danish National Birth Cohort, Preterm delivery study	1748	1.023	30.0 (4.3)	13.3 (1.3)	*"How old were you when you had your first menstrual period?"*
EGCUT	Estonian Genome Center, University of Tartu	987	1.02	41.2 (16.5)	13.4 (1.5)	*"How old you where when you had your first menstruation event?"*
EPIC-obesity cohort	European Prospective Investigation into Cancer and Nutrition - Obesity study cohort	1215	0.97	58.7 (9.0)	12.9 (1.8)	*"How old were you when you had your first menstrual period?"*
EPIC-obesity cases	European Prospective Investigation into Cancer and Nutrition - Obesity study cases	625	0.961	58.8 (8.8)	12.7 (2.0)	*"How old were you when you had your first menstrual period?"*
ERF	Erasmus Rucphen Family study	1103	1.032	47.5 (14.3)	13.1 (1.7)	*"At what age did your menstrual periods begin?"*
FHS	Framingham Heart Study	3801	1.013	42.5 (10.1)	12.8 (1.5)	*"Age at start of menses" and "How old were you when you had your first menstrual period (menses)?" "About how old were you when you had your first menstrual period?"*
HBCS	Helsinki Birth Cohort Study	976	1.006	61.5 (3.0)	12.8 (1.5)	*"At what age did your menstrual periods start?"*
Health 2000 (Genmets) cases	Health2000 cohort - case subsample	457	0.999	51.8 (11.5)	13.4 (1.5)	*"How old were you when your periods started?"*
Health 2000 (Genmets) controls	Health2000 cohort - control subsample	465	1.023	51.9 (11.6)	13.4 (1.6)	*"How old were you when your periods started?"*
InCHIANTI	Invecchiare in Chianti, aging in the Chianti area	597	1.038	68.2 (15.5)	13.3 (1.5)	*"How old were you when you had your first menstrual period?"*
Indiana	Indiana University premenopausal Caucasian women peak BMD study	1497	1.01	33.3 (7.2)	12.6 (1.4)	*At what age did your periods begin? __Years old.*
NFBC	Northern Finland Birth Cohort 1966	2648	1.027	31.2 (0.4)	12.9 (1.3)	*"How old were you when you started menstruating"*
NHS - CGEMS	Nurses' Health Study	2270	1.034	56.8 (6.4)	12.5 (1.4)	*" At what age did your menstrual periods begin?"*
NHS - HU	Nurses' Health Study	3090	1.019	55.7 (6.7)	12.5 (1.4)	*" At what age did your menstrual periods begin?"*
NTR	Netherlands Twin Register	1051	1.01	44.6 (13.6)	13.2 (1.4)	*"How old were you when you had your first menstrual period?"*
QIMR	Queensland Institute of Medical Research	3528	1.029	birth year 1964.8 (19.0)	13.1 (1.3)	*"How old were you when you had your first menstrual period?"*
RS1	Rotterdam Study 1	3175	1.019	69.6 (9.3)	13.5 (1.6)	*"How old were you when you had your first menstrual period?"*
RS2	Rotterdam Study 2	1119	1.002	65.1 (8.4)	13.3 (1.6)	*"How old were you when you had your first menstrual period?"*
RS3	Rotterdam Study 3	1112	1.012	56.2 (6.1)	13.1 (1.6)	*"How old were you when you had your first menstrual period?"*
SAGE	Study of Addiction: Genetics and Environment	1291	1.001	38.4 (9.1)	12.8(1.6)	*"At what age did you have your first menstrual period?"*
SardiNIA	SardiNIA Study	2158	1.225	43.9 (17.2)	13.2 (1.6)	*"At what age did your menstrual periods begin?"*
TwinsUK	TwinsUK	2276	1.006	58.2 (12.7)	13.0 (1.6)	*"How old were you when you had your first menstrual period?"*
TwinsUKII	TwinsUKII	671	1.059	55.4 (14.6)	13.1 (1.6)	*"How old were you when you had your first menstrual period?"*
TwinsUKIII	TwinsUKIII	1016	0.999	62.4 (11.6)	12.9 (1.5)	*"How old were you when you had your first menstrual period?"*
WGHS	Women's Genome Health Study	22028	1.095	54.7 (7.1)	12.4 (1.4)	*"At what age did your menstrual periods begin?'' with response categories ''9 or younger; 10; 11; 12; 13; 14; 15; 16; 17 or older.''*
Overall		87802	1.173			

Supplementary Table 2: Information on genotyping arrays and QC criteria used in each of the Stage 1 discovery studies.

		Genotyping			Imputation and analysis	
Study	**Array**	**Callrate cut-off**	**MAF* cut-off**	**HWE cut-off****	**Imputation program**	**Analysis program**
AGES-Reykjavik	Illumina HumanHap 370K CNV	98%	0.01	1.0E-06	MACH	PLINK
Amish	Affymetrix 500K and 6.0	95%	0.01	1.0E-06	MACH	MMAP***
ARIC	Affymetrix 6.0	90%	0.01	1.0E-06	MACH	ProABEL
B58C-T1DGC	Illumina 550K	NA	0.01	NA	MACH	ProbABEL
B58C-WTCCC	Affymetrix 500K	NA	0.01	NA	IMPUTE	QUICKTEST
CoLaus	Affymetrix 500K	70%	0.01	1.0E-06	IMPUTE	in house Matlab code
deCODE	Illumina HumanHap 300K and 370K CNV	95%	0.01	1.0E-06	IMPUTE	Logistic regression using allele count as a covariate
DNBC	Illumina Human660W-Quad BeadChip	95%	0.01	1.0E-03	MACH	MACH2QTL
EGCUT	Illumina HumanHap 370K CNV	98%	0.01	1.0E-06	IMPUTE	SNPTEST
EPIC-obesity cases	Affymetrix GeneChip 500K	90%	0.01	1.0E-06	IMPUTE	SNPTEST
EPIC-obesity cohort	Affymetrix GeneChip 500K	90%	0.01	1.0E-06	IMPUTE	SNPTEST
ERF	Illumina 6K, 318K, 370K, Afyymetrix 250K	98%	0.01	1.0E-06	MACH	ProbABEL
FHS	Affymetrix 500K + Affymetrix 50K	97%	0.01	1.0E-06	MACH	R-packages
HBCS	Illumina HumanHap610 quad (modified)	95%	0.01	1.0E-06	MACH	ProbABEL
Health 2000 (Genmets) cases	Illumina HumanHap610 quad (modified)	95%	0.01	1.0E-06	MACH	ProbABEL
Health 2000 (Genmets) controls	Illumina HumanHap610 quad (modified)	95%	0.01	1.0E-06	MACH	ProbABEL
InCHIANTI	Illumina HumanHap 550K	98%	0.01	1.0E-04	IMPUTE	SNPTEST
Indiana	Illumina HumanHap 610 Quad version 1B	95%	0.01	1.0E-04	IMPUTE	MERLIN --fastassoc
NHS - CGEMS	Illumina HumanHap 550K	90%	0.01	NA	MACH	ProABEL
NHS - HU	Affymetrix 6.0	98%	0.01	1.0E-04	MACH	ProABEL
NFBC	Illumina Infinium 370CNV Duo	95%	0.01	1.0E-06	MACH	ProbABEL
NTR	Affymetrix 500K Perlegen	95%	0.01	1.0E-05	IMPUTE	SNPTEST
QIMR	Illumina Human610-Quadv1 and 370K CNV	95%	0.01	1.0E-05	MACH	MERLIN --fastassoc
RS1	Illumina HumanHap 550K	98%	0.01	1.0E-06	MACH	MACH2QTL
RS2	Illumina HumanHap 550K	98%	0.01	1.0E-06	MACH	MACH2QTL
RS3	Illumina HumanHap 550K	98%	0.01	1.0E-06	MACH	MACH2QTL
SAGE	Illumina Human 1Mv1_C	98%	0.01	1.0E-04	IMPUTE	SNPTEST
SardiNIA	Affymetrix 10K, 500K	90%	0.05	1.0E-06	MACH	MERLIN --fastassoc
TwinsUK	Illumina HumanHap 300K	95% (MAF 5%) / 99% (MAF 1-5%)	0.01	5.7E-05	IMPUTE	GenABEL
TwinsUKII	Illumina Hap610Quad	95% (MAF 5%) / 99% (MAF 1-5%)	0.01	5.7E-05	IMPUTE	GenABEL
TwinsUKIII	Illumina Hap610Quad	95% (MAF 5%) / 99% (MAF 1-5%)	0.01	5.7E-05	IMPUTE	GenABEL
WGHS	Illumina HumanHap300 Duo "+"	98%	0.01	1.0E-06	MACH	MACH2QTL

*Minor Allele Frequency

**Hardy-Weinburg equilibrium p-value cut-off

***Mixed model analysis for pedigrees

Supplementary Table 3: Results of conditional analyses to verify the presence of additional independent signals for age at menarche.

Supplementary Table 3. Results of conditional analyses to verify the presence of additional independent signals for age at menarche

Possible second signals from the stage 1 meta-analysis

Chr.	Position (B36)	SNP	Alleles[a]	Freq. modelled allele	Effect size (years/allele)	SE	*P*- value[b]	*P*- value[c 2-GC]
2	199561433	rs1947530	A/C	0.66	-0.047	0.008	2.34E-09	3.23E-08
14	99952158	rs10144321	A/G	0.75	0.048	0.008	9.93E-09	1.12E-07

Results from conditional analysis

Chr.	Position (B36)	SNP	Alleles[a]	Freq. coding allele	Effect size (years/allele)	SE	*P*- value[b]	*P*- value[c 2-GC]
2	199561433	rs1947530	A/C	0.66	-0.037	0.008	1.37E-06	8.04E-06
14	99952158	rs10144321	A/G	0.75	0.038	0.009	7.10E-06	3.28E-05

[a]modelled/nonmodelled allele

[b] P-value with genomic control applied to individual studies

[c] P-value with additional adjustment for overall genomic control

(B36) Position according to HapMap Build36

Conditional analyses were adjusted for the top SNP at the 42 genome-wide significant regions (in addition to birth year). Displayed results are based on meta-analysis of all 32 Stage 1 studies.

The possible second signals on chromosomes 2 and 14 failed to reach genome-wide significance in the conditional analyses.

Supplementary Table 4: Details of the 17 studies that contributed to replication for association with age at menarche.

Study	Description	Mean age (SD)	Mean Age at Menarche (SD)	N_SNPs	N women (max)
in silico replication					
BHS	Bogalusa Heart Study	15.8 (4.5)	12.4 (1.2)	42	343
EGCUT	Estonian Genome Center, University of Tartu	38.0 (15.9)	13.3 (1.4)	42	196
INGI - Carlantino	Italian Network of Genetic Isolates	48.1 (19.4)	12.9 (1.6)	42	322
INGI - Friuli Venezia Giulia	Italian Network of Genetic Isolates	50.6 (18.1)	13.2 (1.6)	42	338
INGI - Val Borbera	Italian Network of Genetic Isolates	54.4 (18.3)	12.9 (1.5)	42	910
KORA F3	Cooperative Health Research in the Region of Augsburg, KOoperative Gesundheitsforschung in der Region Augsburg	61.8 (10.1)	13.7 (1.5)	41	809
KORA S4	Cooperative Health Research in the Region of Augsburg, KOoperative Gesundheitsforschung in der Region Augsburg	53.5 (8.8)	13.5 (1.5)	42	898
SASBAC cases	Singapore and Swedish Breast Cancer Study	62.6 (6.3)	13.4 (1.4)	41	723
SASBAC controls	Singapore and Swedish Breast Cancer Study	62.6 (6.3)	13.5 (1.4)	40	685
SEARCH	Ovarian cancer cases	birth year 1942.7 (9.9)	12.8 (1.5)	42	1126
STR_MZ twins*	Swedish National Twin Cohort	65.3 (5.82)	13.7 (1.28)	41	151
Raine	Raine Study, Western Australia	14.1 (0.19)**	12.8 (1.13)	42	527
Orcades	Orkney Complex Disease Study, EUROpean Special Populations reseArch Network	52.7 (15.3)	12.8 (1.4)	42	348
SPLIT	Split, Croatia	46.7 (13.9)	13.5 (1.6)	42	283
KORCULA	Korcula Island, Croatia	54.8 (13.5)	13.5 (1.6)	42	508
VIS	Vis Island, Croatia (EUROSPAN)	56.2 (13.5)	13.5 (1.7)	42	502
de novo replication					
ALSPAC mothers	Avon Longitudinal Study of Parents and Children	28.2 (4.8)	12.8 (1.5)	30	6,118

*302 twin pairs with average age at menarche used

**Girls not reaching menarche by this exam were asked to return a slip of paper reporting the dates of their first two periods to the investigators

Supplementary Table 5: Meta-analysed results of replication studies (up to 14,731 women).

SNP	Nearest gene(s)	MAF[a]	Alleles[b]	N	Beta[c]	SE	*P*- value
Previous menarche loci							
rs7759938	*LIN28B*	0.32	C/T	14,185	6.3	1.0	4.6E-11
rs2090409	*TMEM38B*	0.31	A/C	14,708	-4.4	0.9	2.7E-06
30 novel menarche loci							
rs1079866	*INHBA*	0.15	G/C	14731	1.7	1.3	1.9E-01
rs466639	*RXRG*	0.13	T/C	14279	-2.9	1.3	3.1E-02
rs6438424	*3q13.32*	0.50	A/C	8634	-3.0	1.1	6.7E-03
rs1398217	*FUSSEL18*	0.43	G/C	14344	-2.7	0.9	2.3E-03
rs12617311	*PLCL1*	0.32	A/G	14007	-2.5	1.0	1.1E-02
rs9635759	*CA10*	0.32	A/G	14002	2.6	1.0	1.1E-02
rs6589964	*BSX*	0.48	A/C	13754	-1.6	0.9	8.3E-02
rs10980926	*ZNF483*	0.36	A/G	14227	0.8	0.9	3.8E-01
rs17268785	*CCDC85A*	0.17	G/A	14233	2.9	1.2	1.5E-02
rs13187289	*PHF15*	0.20	G/C	14303	2.8	1.2	1.4E-02
rs7642134	*VGLL3*	0.38	A/G	14205	-2.8	0.9	2.1E-03
rs17188434	*NR4A2*	0.07	C/T	14356	-2.2	1.8	2.2E-01
rs2002675	*TRA2B, ETV5*	0.42	G/A	14334	2.5	0.9	6.6E-03
rs7821178	*PXMP3*	0.34	A/C	14151	-1.7	0.9	8.0E-02
rs1659127	*MKL2*	0.34	A/G	14021	2.2	1.0	2.5E-02
rs10423674	*CRTC1*	0.35	A/C	13543	1.6	1.0	1.1E-01
rs10899489	*GAB2*	0.15	A/C	14201	1.4	1.2	2.5E-01
rs6575793	*BEGAIN*	0.42	C/T	13899	0.7	1.0	4.6E-01
rs4929923	*TRIM66*	0.36	T/C	8510	2.9	1.2	1.6E-02
rs6439371	*TMEM108, NPHP3*	0.34	G/A	8581	2.6	1.2	3.0E-02
rs900145	*ARNTL*	0.30	C/T	8649	2.3	1.2	6.5E-02
rs6762477	*RBM6*	0.44	G/A	12447	1.4	1.0	1.5E-01
rs2947411	*TMEM18*	0.17	A/G	8657	3.4	1.4	1.9E-02
rs1361108	*C6orf173, TRMT11*	0.46	T/C	14126	-1.7	0.9	6.0E-02
rs1364063	*NFAT5*	0.43	C/T	8669	3.0	1.1	7.1E-03
rs633715	*SEC16B*	0.20	C/T	14274	-1.5	1.1	1.9E-01
rs4840086	*PRDM13, MCHR2*	0.42	G/A	8669	-2.0	1.1	7.5E-02
rs7617480	*KLHDC8B*	0.22	A/C	14341	1.2	1.0	2.4E-01
rs9939609	*FTO*	0.40	A/T	8665	0.7	1.2	5.3E-01
rs852069	*PCSK2*	0.37	A/G	14306	-0.9	0.9	3.3E-01
10 possible menarche loci							
rs757647	*KDM3B*	0.22	A/G	14326	-0.8	1.1	4.4E-01
rs9555810	*C13orf16, ARHGEF7*	0.28	G/C	14266	0.7	1.0	4.9E-01
rs16938437	*PHF21A*	0.09	T/C	14330	-1.4	1.6	3.8E-01
rs2687729	*EEFSEC*	0.27	G/A	8669	1.3	1.3	3.2E-01
rs1862471	*OLFM2*	0.47	G/C	13470	-0.1	1.0	9.4E-01
rs12472911	*LRP1B*	0.20	C/T	8585	2.1	1.4	1.4E-01
rs3914188	*ECE2*	0.27	G/C	14085	-0.3	1.0	7.9E-01
rs2243803	*SLC14A2*	0.40	A/T	8659	1.0	1.1	3.9E-01
rs3743266	*RORA*	0.32	C/T	8666	-0.3	1.2	7.8E-01
rs7359257	*IQCH*	0.45	A/C	14303	-0.5	0.9	6.0E-01

[a] Minor allele frequency

[b] Minor / Major allele

[c] Per allele change in age at menarche (weeks)

Supplementary Table 6: Variance in age at menarche explained by the 42 known, confirmed and possible novel loci in each of the 17 replication studies.

Study	Description	N women in full model	N_SNPs	Variance Explained (r2)
in silico replication				
SEARCH	Ovarian cancer cases	1126	42	0.036
INGI - Val Borbera	Italian Network of Genetic Isolates	910	42	0.061
KORA S4	Cooperative Health Research in the Region of Augsburg, KOoperative Gesundheitsforschung in der Region Augsburg	898	42	0.054
KORA F3	Cooperative Health Research in the Region of Augsburg, KOoperative Gesundheitsforschung in der Region Augsburg	805	41	0.058
KORCULA	Korcula Island, Croatia (EUROSPAN)	508	41	0.077
VIS	Vis Island, Croatia (EUROSPAN)	502	42	0.150
Orcades (Eurospan)	Orkney Complex Disease Study, EUROpean Special Populations reseArch Network (EUROSPAN)	348	40	0.110
BHS	Bogalusa Heart Study	338	42	0.099
INGI - Friuli Venezia Giulia	Italian Network of Genetic Isolates	338	42	0.151
INGI - Carlantino	Italian Network of Genetic Isolates	322	42	0.181
SPLIT	Split, Croatia	283	42	0.163
EGCUT	Estonian Genome Center, University of Tartu	196	42	0.186
SASBAC controls	Singapore and Swedish Breast Cancer Study	188	42	0.238
SASBAC cases	Singapore and Swedish Breast Cancer Study	158	41	0.256
STR_MZ twins	Swedish National Twin Cohort	151	42	0.158
Raine	Raine Study, Western Australia	88	42	0.454
de novo replication				
ALSPAC mothers	Avon Longitudinal Study of Parents and Children	3938	30	0.011

Sorted by sample size. The largest four in silico replication studies are highlighted as giving rise to more reliable estimates.

Supplementary Table 7: Biological functions of genes at/near the 42 known, confirmed or possible novel menarche loci.

SNP	Nearest gene(s)	Distance from gene (kb)	Chr	Gene name & function
Previous menarche loci				
rs7759938	*LIN28B*	~26kb	6	**lin-28 homolog B (C. elegans).** Also associated with short adult height. Regulates let-7 micro-RNA processing. Deletarious mutations in lin-28 in C.
rs2090409	*TMEM38B*	~400kb	9	**transmembrane protein 38B.** Trimeric intracellular cation channel type B.
30 novel menarche loci				
rs1079866	*INHBA*	~250kb	7	**inhibin, beta A.** Joins the alpha subunit to form an inhibitor of pituitary secretion of follicle stimulating hormone (FSH). Circulating levels of inhibin A rise with gonadal growth during puberty and negatively feedback on the pituitary.
rs466639	*RXRG*	intronic	1	**retinoid X receptor, gamma.** A nuclear receptor which forms dimers with the receptors for retinoic acid, thyroid hormone, and vitamin D, increasing both DNA binding and transcriptional function on their respective response elements.
rs6438424	*3q13.32*	*intergenic*	3	
rs1398217	*FUSSEL18*	intronic	18	**functional smad suppressing element 18.** Specifically neuronally-expressed member of a family of nuclear oncoproteins that repress transforming growth factor-beta (TGF-beta) signaling through inhibition of transcriptional activity of Smad proteins.
rs12617311	*PLCL1*	~195kb	2	**phospholipase c-like 1.** Involved in an inositol phospholipid-based intracellular signaling cascade.
rs9635759	*CA10*	~94kb	17	**carbonic anhydrase X.** Encodes a protein with likely no carbonic anhydrase catalytic activity, but is expressed widely in the central nervous system.
rs6589964	*BSX*	~18kb	11	**brain specific homebox.** A DNA binding protein and transcriptional activator. In the mouse Bsx is expressed specifically in the pineal gland, telencephalic septum, hypothalamic pre-mammillary body and arcuate nucleus, and is necessary for postnatal growth.
rs10980926	*ZNF483*	intronic	9	**zinc finger protein 483.** Unknown function. Is conserved in chimpanzee, dog, mouse, and rat.
rs17268785	*CCDC85A*	intronic	2	**coiled-coil domain containing 85A.** Unknown function.
rs13187289	*PHF15*	~12kb	5	**PHD finger protein 15.** Encodes a PHD zinc finger protein transcription factor. Has been reported to be downregulated in human peripheral blood mononuclear cells by melatonin.
rs7642134	*VGLL3*	~70kb	3	**vestigial like 3 (Drosophila).** A transcriptional regulator with a possible role in the suppression of ovarian tumorigenesis.
rs17188434	*NR4A2*	~84kb	2	**nuclear receptor subfamily 4, group A, member 2.** Encodes a transcription factor essential for the differentiation of dopaminergic neurons in substantia nigra. Mutations are associated with disorders of dopaminergic dysfunction, including Parkinsons disease.
rs2002675	*TRA2B; ETV5*	~4kb, ~135kb	3	**transformer 2 beta homolog (Drosophila).** Related to the ancestral gene in Drosophila, Transformer-2 (Tra2), which encodes an RNA-binding protein, an important regulator of sex determination. **ets variant 5.** A member of the ets family of transcription factors.
rs7821178	*PXMP3*	~181kb	8	**peroxisomal membrane protein 3.** Encodes an integral peroxisomal membrane protein required for peroxisome biogenesis. Mutations result in the peroxisomal biogenesis disorders Zellweger syndrome and infantile Refsum disease.
rs1659127	*MKL2*	~28kb	16	**MKL/myocardin-like 2**. A broadly expressed member of a family of transcriptional coactivators which bind to serum response factor (SRF) and strongly activate transcription from promoters with SRF binding sites.
rs10423674	*CRTC1*	intronic	19	**CREB regulated transcription coactivator 1.** Encodes an activator of cellular gene expression. Crtc1(-/-) mice are hyperphagic, obese and infertile, and females have low circulating luteinizing hormone levels. Leptin potentiates the effects of Crtc1 transcriptional activity, and Crtc1 over-expression in hypothalamic cells increases Kiss1 gene expression.
rs10899489	*GAB2*	intronic	11	**GRB2-associated binding protein 2.** Encodes the principal activator of phosphatidylinositol-3 kinase in response to activation of the high affinity IgE receptor. Forms a complex with growth factor receptor-bound protein 2 (Grb2) involved in signal transduction/cell communication. Has been associated with Alzheimer's disease and implicated in angiogenesis and melanoma tumor progression and metastasis.
rs6575793	*BEGAIN*	intronic	14	**brain-enriched guanylate kinase-associated homolog (rat).** Lies in the Dlk1 / Gtl2 imprinting region and shows transcript-specific genomic imprinting in the mouse (preferential paternal transcription).
rs4929923	*TRIM66*	3'UTR	11	**tripartite motif-containing 66.** Encodes TIF1delta, one of the Transcriptional intermediary factor 1 (TIF1) proteins which are encoded by a family of highly conserved developmental and physiological control genes. These proteins have been implicated in epigenetic mechanisms of transcriptional repression. Uniquely, TRIM66 expression is confined to haploid elongating spermatids.
rs6439371	*TMEM108; NPHP3*	~146kb, ~170kb	3	**transmembrane protein 108**. Unknown function. **nephronophthisis 3 (adolescent)**. Possible role in renal tubular development and function.
rs900145	*ARNTL*	~5 kb	11	**aryl hydrocarbon receptor nuclear translocator-like** (Also known as *BMAL1*). The encoded protein forms a complex with CLOCK to activate PER1 and possibly other circadian rhythm-associated genes. Expression in the ovary is elevated following the LH surge. *Bmal1* null mice lack central and peripheral cellular rhythms and have delayed puberty.
rs6762477	*RBM6*	intronic	3	**RNA binding motif protein 6.** An RNA binding protein which potentially modulates apoptosis and has been implicated in myeloid proliferative disease.
rs2947411	*TMEM18*	~53kb	2	**transmembrane protein 18.** Modulates the migration of neural stem cells.
rs1361108	*C6orf173; TRMT11*	~98kb, ~407kb	6	**chromosome 6 open reading frame 173**. Is commonly up-regulated in many human cancer tissues and possesses transforming activities in mouse fibroblast cell line. **tRNA methyltransferase 11 homolog (S. cerevisiae)**. Putative member of the RNA methylase enzyme family.
rs1364063	*NFAT5*	~10kb	16	**nuclear factor of activated T-cells 5, tonicity-responsive**. Encodes a member of the nuclear factors of activated T cells family of transcription factors, which regulate the inducible gene transcription during the immune response.
rs633715	*SEC16B*	~44kb	1	**SEC16 homolog B (S. cerevisiae).** Homolog of an ancestral gene that is required for the organization of transitional endoplasmic reticulum (ER) sites and protein export.
rs4840086	*PRDM13, MCHR2*	~145kb, ~160kb	6	**PR domain containing 13.** Regulation of transcription. **melanin concentrating hormone receptor 2.** This orphan G protein-coupled receptor shows high affinity binding to the neuropeptide melanin-concentrating hormone (MCH), which is known to regulate energy homeostasis and mood via MCHR1.
rs7617480	*KLHDC8B*	intronic	3	**lysine (K)-specific demethylase 3B.** Is a candidate tumor suppressor gene based on its predicted function, a regulator of chromatin remodeling, and
rs9939609	*FTO*	intronic	16	**fat mass and obesity associated.** Involved in nucleic acid demethylation, but exact function is unknown. Is transcriptionally regulated by feeding and fasting.
rs852069	*PCSK2*	~84kb	20	**proprotein convertase subtilisin/kexin type 2.** Cleaves latent precursor proteins, such as proinsulin and proopiomelanocortin, into their biologically active products. PCSK2 differs from PCSK1 in that it additionally cleaves proluteinizing-hormone-releasing hormone.
10 possible menarche loci[j]				
rs757647	*KDM3B*	intronic	5	**lysine (K)-specific demethylase 3B.** Is a candidate tumor suppressor gene based on its predicted function, a regulator of chromatin remodeling, and clonogenic growth suppressive activities.
rs9555810	*C13orf16, ARHGEF7*	~185kb, ~223kb	13	**chromosome 13 open reading frame 16.** Uncharacterised protein. **Rho guanine nucleotide exchange factor (GEF) 7**. Belongs to a family of cytoplasmic proteins that activate Rho proteins by exchanging bound GDP for GTP. ARHGEF7 has a role in cell proliferation through phosphorylation of FOXO3a. FOXO3a knockout mice are infertile due to early depletion of the follicle pool. Patients with microdeletions of 13q33-34 have mental retardation and microcephaly, and male patients frequently have genital malformations.
rs16938437	*PHF21A*	intronic	11	**PHD finger protein 21A.** A component of the BRAF-HDAC complex (BHC), PHF21A modulates the repression of neurosecretion through the transcription repressor element-1 silencing transcription factor (REST). Possible role in spermatogenesis.
rs2687729	*EEFSEC*	intronic	3	**eukaryotic elongation factor, selenocysteine-tRNA-specific**. Regulates selenocysteine incorporation into mRNAs in the generation of selenoproteins.
rs1862471	*OLFM2*	intronic	19	**olfactomedin 2.** Is one of a diverse family of olfactomedins, which are secreted glycoproteins that have been implicated in the timing and growth of the neural crest, and include the glaucoma-associated protein myocilin.
rs12472911	*LRP1B*	intronic	2	**low density lipoprotein-related protein 1B.** Belongs to the low density lipoprotein (LDL) receptor gene family, which have a wide variety of roles in normal cell function and development. LRP1B has been implicated in tumorigenesis and associated with ageing.
rs3914188	*ECE2*	3'UTR	3	**endothelin converting enzyme 2.** A type II metalloprotease that generate functionally pleiotropic members of the endothelin vasoactive peptide family. Downregulated in Alzheimer's disease.
rs2243803	*SLC14A2*	~238kb	18	**solute carrier family 14 (urea transporter), member 2 (SLC14A2)**. Encodes the renal tubular transporter of urea.
rs3743266	*RORA*	3'UTR	15	**RAR-related orphan receptor A.** Encodes the retinoid-related orphan receptor, RORalpha, a nuclear hormone receptor critical for the development of the cerebellum. May play a role in maturation of photoreceptors and regulation of the circadian clock.
rs7359257	*IQCH*	intronic	15	**IQ motif containing H.** Unknown function.

Supplementary Table 8: Functional networks and their relevant functions for 42 candidate genes associated with age at menarche.

	Age at menarche associated genes in Network	Other related genes/molecules in network	Associated network functions
Network 1: $p=1x10^{-37}$	16 genes: ARHGEF7, CRTC1, ECE2, ETV5, INHBA, KLHDC8B, MKL2, NFAT5, OLFM2, PHF15, PHF21A, PLCL1, RBM6, SEC16B, SLC14A2, TRA2B	AR, CRYAB, E2F1, FABP4, HNF4A, INHBC, INHBE, MIR292, MIR136 (includes EG:406927), MIRLET7B (includes EG:406884), MYC, NBAS, NCOA3, PCK1, RARG, RXRB, SNAP25, ST13, YWHAG	Gene Expression, Cellular Growth and Proliferation, Cellular Function and Maintenance
Network 2: $p=1x\ 10^{-23}$	11 genes: ARNTL, BEGAIN, CCDC85A, GAB2, LRP1B, MCHR2, NR4A2, PCSK2, PXMP3, RORA, RXRG	ACADM, Acox, ACSL1, BGLAP, CYP24A1, CYP4A11, DLG4, EHHADH, FABP, FABP1, FABP3, FABP5, HMG CoA synthase, ITGB3BP, LIVTR, MIR24-1 (includes EG:407012), MMP11, PMCH, RBP2, RQCD1, RXRA, SRC, TLE1, ZFP64	Lipid Metabolism, Small Molecule Biochemistry, Molecular Transport

		p-value	***# focus genes***
Diseases and disorders	Genetic disorder	1.0E-03 - 3.5E-02	24
	Neurological disease	1.0E-03 - 2.7E-02	17
	Psychological disorders	1.0E-03 - 7.0E-03	10
	Endocrine system disorders	1.4E-03 - 1.6E-02	13
	Metabolic disease	1.4E-03 - 1.6E-02	14
Molecular and cellular functions	Gene expression	1.5E-05 - 4.4E-02	14
	Small molecule biochemistry	1.4E-04 - 4.8E-02	7
	Cellular development	5.1E-04 - 4.8E-02	8
	Cellular growth and proliferation	5.1E-04 - 4.8E-02	7
	Antigen presentation	2.3E-03 - 2.3E-03	1

Ingenuity Pathway Analysis (IPA) Knowledge Base 8.5 (Ingenuity Systems, CA, USA) was used to explore the functional relationship between proteins encoded by the genes or nearest genes of the 42 genome-wide significant loci. The genes were entered into the Ingenuity database and analyzed for direct interactions only. Networks were generated with a maximum size of 35 genes.

Supplementary Table 9: Pathways significantly associated with age at menarche identified by GSEA.

			95th percentile enrichment cutoff				75th percentile enrichment cutoff				
Database	**Gene set**	**# genes in category tested by GSEA**	**Nominal GSEA P-value**	**FDR corrected P-value**	**Observed # genes above cutoff**	**Expected # genes above cutoff**	**Nominal GSEA P-value**	**FDR corrected P-value**	**Observed # genes above cutoff**	**Expected # genes above cutoff**	**Genes nearest to validated age at menarche SNPs**
Panther	Coenzyme A biosynthesis	7	2.0E-04	4.9E-03	4	0	1.2E-02	1.4E-01	5	2	
Panther	Angiotensin II stimulated signaling through G proteins and beta-arrestin	5	6.0E-04	6.2E-03	3	0	1.0E-01	3.9E-01	3	1	
PANTHER BIOLOGICAL PROCESS	Stress response	172	4.7E-02	9.0E-01	14	9	3.7E-05	5.1E-03	67	43	NR4A2, NFAT5
PANTHER BIOLOGICAL PROCESS	Coenzyme metabolism	52	7.0E-04	1.5E-01	9	3	3.0E-04	1.9E-02	25	13	
PANTHER BIOLOGICAL PROCESS	General mRNA transcription activities	39	5.9E-01	8.9E-01	2	2	4.0E-04	4.4E-02	19	10	

MAGENTA (Ref 44) was used to test for enrichment of multiple modest associations with age at menarche in 2529 pathways from Gene Ontology, PANTHER, KEGG and Ingenuity.
Only gene sets with FDR<0.05 are presented here. GSEA P-values in bold refer to P-values that pass the Bonferroni correction cutoff (accounting for testing two enrichment cutoffs).

Supplementary Table 10: The 42 known, confirmed or possible novel menarche loci related to lymphoblastoid cell line (LCL) eQTLs.

Nearest gene(s)	SNP	chr	position	Gene Expressed	Mechanism	P-value
LIN28B	rs7759938	6	105485647	-	-	-
TMEM38B	rs2090409	9	108006909	-	-	-
INHBA	rs1079866	7	41436618	-	-	-
RXRG	rs466639	1	163661506	*TEX12*	Trans (Chr 11)	2x10-7
3q13.32	rs6438424	3	119057512	-	-	-
BSX	rs6589964	11	122375893	-	-	-
CA10	rs9635759	17	46968784	-	-	-
ZNF483	rs10980926	9	113333455	-	-	-
PLCL1	rs12617311	2	199340810	-	-	-
FUSSEL18	rs1398217	18	43006236	-	-	-
FTO	rs9939609	16	52378028	-	-	-
NR4A2	rs17188434	2	156805022	-	-	-
CCDC85A	rs17268785	2	56445587	-	-	-
BEGAIN	rs6575793	14	100101970	-	-	-
PHF15	rs13187289	5	133877076	-	-	-
GAB2, ZNF75C	rs10899489	11	77773021	*GAB2*	Cis	3x10-7
OLFM2	rs1862471	19	9861322	-	-	-
C13orf16, ARHGEF7	rs9555810	13	110979438	-	-	-
PXMP3	rs7821178	8	78256392	-	-	-
CRTC1	rs10423674	19	18678903	-	-	-
PCSK2	rs852069	20	17070593	-	-	-
KDM3B	rs757647	5	137735214	*ATP6V1H*	Trans (Chr 8)	6x10-7
PHF21A	rs16938437	11	46009151	-	-	-
RBM6	rs6762477	3	50068213	*RBM6*	Cis	4.8x10-11
SEC16B	rs633715	1	176119203	-	-	-
KLHDC8B	rs7617480	3	49185736	-	-	-
VGLL3	rs7642134	3	86999572	-	-	-
ECE2	rs3914188	3	185492742	-	-	-
C6orf173, TRMT11	rs1361108	6	126809293	-	-	-
MKL2	rs1659127	16	14295806	-	-	-
TRA2B, ETV5	rs2002675	3	187112262	-	-	-
IQCH	rs7359257	15	65489961	-	-	-
ARNTL	rs900145	11	13250481	-	-	-
PRDM13, MCHR2	rs4840086	6	100315159	-	-	-
EEFSEC	rs2687729	3	129377916	-	-	-
TMEM108, NPHP3	rs6439371	3	134093442	-	-	-
TMEM18	rs2947411	2	604168	-	-	-
TRIM66	rs4929923	11	8595776	-	-	-
RORA	rs3743266	15	58568805	*NARG2*	Cis	7x10-7
SLC14A2	rs2243803	18	41210670	-	-	-
LRP1B	rs12472911	2	141944979	-	-	-
NFAT5	rs1364063	16	68146073	-	-	-

Using publicly available LCL dataset (http://www.sph.umich.edu/csg/liang/asthma/ mRNA by SNP Browser) from Dixon et al, Nature Genetics 2007. Only P-values <1.0E7 are shown.

Supplementary Table 11: Adipose tissue eSNPs related to age at menarche.

eSNP	Chr	Position	Menarche P-value	eSNP P-value	Transcript	Closest gene	Distance	Allele 1	Allele 2	Allele 1 frequency
rs4660740	1	43921204	4.6E-05	1.2E-06	MED8	*JMJD2A*	22571	a	g	0.744
rs823096	1	203946510	8.9E-05	1.5E-08	PM20D	*NUCKS1*	7221	t	g	0.439
rs823114	1	203986155	1.2E-04	8.0E-13	PM20D	*NUCKS1*	239	a	g	0.540
rs1378410	2	199182802	1.2E-04	4.4E-05	MARS2	*PLCL1*	461509	a	g	0.248
rs4955439	3	49220649	**2.8E-08**	2.5E-14	WDR6	*CCDC36*	9785	t	g	0.262
rs4955417	3	49273209	**5.5E-08**	7.3E-09	WDR6	*CCDC36*	3050	t	c	0.262
rs2236950	3	50395558	2.2E-05	8.9E-16	HYAL3	*CACNA2D2*	20323	a	c	0.184
rs9310074	3	88227760	2.4E-06	5.6E-05	ZNF654	*CGGBP1*	36924	t	c	0.843
rs4687889	3	119020129	**1.6E-11**	1.3E-04	AK022000	*IGSF11*	1082041	t	c	0.532
rs10512868	3	134067262	1.9E-05	3.5E-06	NPHP3	*NPHP3*	143296	a	g	0.173
rs4912539	3	185544081	4.1E-06	6.8E-05	EIF2B5	*FAM131A*	2675	a	g	0.700
rs4698894	4	104540550	1.2E-05	3.9E-08	CISD2	*TACR3*	189523	a	g	0.760
rs329319	5	133934508	3.6E-05	3.8E-05	PITX1	*PHF15*	12309	a	g	0.428
rs2524044	6	31364732	7.9E-06	1.6E-06	HCG18	*HLA-C*	16898	t	g	0.834
rs222461	6	52997234	1.8E-04	6.3E-09	FBXO9	*ICK*	23178	a	c	0.752
rs11188661	10	97950983	1.2E-04	1.6E-07	AK000974	*BLNK*	9531	a	g	0.314
rs10501087	11	27626684	1.2E-04	1.4E-04	LIN7C	*BDNF*	6333	t	c	0.798
rs10835211	11	27657941	9.4E-06	9.2E-12	LIN7C	*BDNF*	19815	a	g	0.252
rs4944196	11	77686379	**8.3E-09**	9.3E-05	C11orf67	*GAB2*	44195	a	g	0.160
rs7160413	14	100185286	2.1E-05	1.5E-04	C14orf134	*DLK1*	77719	a	g	0.079
rs4782294	16	19774522	8.0E-05	2.0E-06	GPRC5B	*IQCK*	1836	t	c	0.128
rs1398883	17	39115918	1.5E-04	2.0E-04	KRTHA8	*MEOX1*	21130	a	g	0.408
rs2836961	21	39548890	8.1E-06	4.0E-07	WRB	*BRWD1*	41345	a	c	0.612

Menarche P-values are derived from our Stage 1 meta-analysis of 32 studies

All 23 eSNPs were significantly associated with age at menarche after correction for multiple testing (1/n threshold for 5,184 independent tests was P<1.93 x 10-4).

eSNP P-values are derived from the Icelandic Family Adipose cohort (Ref 26)

Supplementary Table 12: Age at menarche associations for 8,770 SNPs in 16 candidate genes and their surrounding regions (+/-300kb).

See Excel file online.

Supplementary Table 13: Associations between known obesity-related SNPs and age at menarche.

Nearby Gene	SNP*	Chr	Obesity Phenotype	Menarche Beta (weeks/allele)	Menarche SE	Menarche *P* value	Obesity-susceptibility allele	Menarche-decreasing allele
FTO	rs9939609	16q12	BMI	2.5	0.4	**3.3E-11**	A	A
SEC16B	rs10913469	1q25	BMI	2.6	0.5	**2.4E-08**	C	C
GNPDA2	rs10938397	4p13	BMI	2.1	0.4	**8.7E-08**	G	G
NEGR1	rs2815752	1p31	BMI	1.9	0.4	**5.9E-07**	A	A
TMEM18	rs6548238	2p25	BMI	2.7	0.5	**7.1E-07**	C	C
FAIM2	rs7138803	12q13	BMI	1.8	0.4	**1.7E-06**	A	A
BDNF	rs4923461	11p14	BMI	1.7	0.5	**3.1E-04**	A	A
KCTD15	rs11084753	19q13	BMI	1.4	0.4	**5.9E-04**	G	G
TRA2B, ETV5	rs7647305	3q27	BMI	1.2	0.5	**9.0E-03**	C	C
MTCH2	rs10838738	11p11	BMI	0.6	0.4	1.4E-01	G	G
MC4R	rs17782313	18q21	BMI	0.6	0.4	1.5E-01	C	T
SH2B1	rs7498665	16p11	BMI	0.2	0.4	5.7E-01	G	G
TFAP2B	rs987237	6p12	WHR	1.6	0.5	**7.8E-04**	G	G
MSRA	rs7826222	8p23	WHR	1.8	0.8	**2.4E-02**	G	G
NRXN3	rs10146997	14q31	WHR	0.7	0.5	1.4E-01	G	G
LYPLAL1	rs2605100	1q41	WHR	0.3	0.4	4.6E-01	G	G
NPC1	rs1805081	18q11	Obesity	0.7	0.4	5.1E-02	T	T
PTER	rs10508503	10p12	Obesity	1.1	0.7	1.0E-01	C	T
MAF	rs1424233	16q23	Obesity	0.1	0.4	8.3E-01	T	C

WHR: waist-hip ratio

Menarche P-values are derived from our Stage 1 meta-analysis of 32 studies with genomic control applied to individual studies.

*Selected SNPs at each locus are those published for association with BMI/WHR/obesity (rather than those with the strongest signal for age at menarche)

Supplementary Table 14: Associations between known height SNPs and age at menarche.

Gene	SNP*	Chr	Position	Menarche Beta (weeks/allele)	Menarche SE	Menarche *P* value	Height-increasing allele	Menarche-increasing allele
LIN28B	rs314277	6	105514355	6.9	0.6	**2.1E-35**	a	a
PXMP3	rs7846385	8	78322734	2.5	0.4	**1.9E-09**	***c***	***t***
C6orf173	rs4549631	6	127008001	1.8	0.4	**4.9E-07**	***c***	***t***
SCMH1	rs6686842	1	41303458	-1.1	0.4	**3.3E-03**	***t***	***c***
Histone cluster 1	rs10946808	6	26341366	1.1	0.4	**6.4E-03**	a	a
NOG	rs4794665	17	52205328	-0.9	0.4	**1.1E-02**	***a***	***g***
HMGA2	rs1042725	12	64644614	-0.8	0.4	**2.0E-02**	c	c
TBX2	rs757608	17	56852059	-0.9	0.4	**2.2E-02**	***a***	***g***
HLA Class III	rs2844479	6	31680935	-0.9	0.4	**2.4E-02**	***a***	***c***
ZBTB38	rs6440003	3	142576899	0.8	0.4	**3.5E-02**	a	a
CABLES1	rs4800148	18	18978326	-1.0	0.5	**3.7E-02**	***a***	***g***
ZNF462	rs4743034	9	108672174	0.8	0.4	6.4E-02	a	a
PPARD	rs2814993	6	34726871	0.9	0.5	7.2E-02	a	a
PPARD	rs4713858	6	35510763	-0.9	0.5	8.7E-02	g	g
CDK6	rs2282978	7	92102346	-0.6	0.4	1.3E-01	c	c
Histone cluster 2	rs11205277	1	148159496	0.6	0.4	1.4E-01	g	a
PTCH1	rs10512248	9	97299524	-0.6	0.4	1.5E-01	g	g
SPAG17	rs12735613	1	118685496	0.6	0.4	2.0E-01	g	a
HLA Class III	rs185819	6	32158045	-0.5	0.4	2.0E-01	t	c
HMGA1	rs1776897	6	34302989	1.2	0.9	2.0E-01	g	t
DYM	rs8099594	18	45245158	0.5	0.4	2.5E-01	a	a
RNF135	rs3760318	17	26271841	0.4	0.4	2.6E-01	g	a
NCAPG	rs16896068	4	17553938	-0.6	0.5	2.8E-01	g	g
DLEU7	rs3116602	13	50009356	-0.5	0.5	3.0E-01	t	g
GPR126	rs4896582	6	142745570	-0.4	0.4	3.0E-01	g	g
HHIP	rs1812175	4	145794294	-0.5	0.5	3.1E-01	g	g
TSEN15	rs2274432	1	182287568	-0.4	0.4	3.2E-01	a	g
FBLN5	rs8007661	14	91529711	0.4	0.5	4.4E-01	c	t
ANAPC13	rs10935120	3	135715782	-0.3	0.4	4.6E-01	g	g
BMP6	rs12198986	6	7665058	0.3	0.4	4.9E-01	a	a
GDF5	rs6060369	20	33370575	0.2	0.4	5.4E-01	c	t
ADAMSTSL3	rs2562784	15	82077496	0.3	0.5	5.6E-01	g	a
DOT1L	rs12986413	19	2121954	-0.2	0.4	6.2E-01	t	t
PLAG1	rs10958476	8	57258362	-0.2	0.5	6.9E-01	c	c
ADAMSTS17	rs4533267	15	98603794	-0.2	0.4	7.0E-01	a	g
EFEMP1	rs3791679	2	55950396	0.2	0.4	7.3E-01	a	a
ZNF678	rs1390401	1	225864573	0.2	0.5	7.5E-01	a	a
DNM3	rs678962	1	170456512	0.1	0.5	7.5E-01	g	t
SOCS2	rs11107116	12	92502635	0.1	0.4	7.6E-01	t	t
PLAG1	rs9650315	8	57318152	-0.2	0.6	7.9E-01	g	g
BMP2	rs967417	20	6568893	0.1	0.4	8.4E-01	g	a
GNA12	rs798544	7	2729628	-0.1	0.4	8.5E-01	c	c
IHH	rs6724465	2	219652090	-0.1	0.6	9.3E-01	g	g
ACAN	rs8041863	15	87160693	0.0	0.4	9.7E-01	a	a

Chi-square =7.02, P=0.008 for 11/44 SNPs associated with age at menarche (at P<0.05) vs. 2.2 expected by chance

However 7 height-increasing SNPs are associated with earlier menarche, and 4 with later menarche.

Menarche P-values are derived from our Stage 1 meta-analysis of 32 studies with genomic control applied to individual studies

Supplementary Table 15: Association of menarche loci with BMI in up to 32,530 adults in the GIANT consortium.

								Association with menarche				Association with BMI		
SNP	Nearest gene(s)	Distance from gene (kb)	Chr	Position (B36)	MAF	Minor allele	Major allele	Menarche decreasing allele	Beta (weeks/allele)	se	P-value	BMI increasing allele	P-value	Direction consistent[a]
rs9939609	FTO	intronic	16	52378028	0.40	A	T	A	2.12	0.39	3.1E-08	A	6.3E-17	y
rs2947411	TMEM18	~53kb	2	604168	0.17	A	G	G	2.84	0.51	1.7E-08	G	3.0E-05	y
rs4929923	TRIM66	3'UTR	11	8595776	0.36	T	C	C	2.27	0.40	1.2E-08	C	2.3E-03	y
rs3914188	ECE2	3'UTR	3	185492742	0.27	G	C	G	2.21	0.43	2.6E-07	G	5.6E-03	y
rs633715	SEC16B	~44kb	1	176119203	0.20	C	T	C	2.63	0.47	2.1E-08	C	6.8E-03	y
rs10899489	GAB2	intronic	11	77773021	0.15	A	C	C	3.09	0.54	8.1E-09	C	1.0E-02	y
rs466639	RXRG	intronic	1	163661506	0.13	T	C	T	4.20	0.57	1.3E-13	T	1.1E-02	y
rs13187289	PHF15	~12kb	5	133877076	0.20	G	C	C	3.02	0.48	1.9E-10	C	2.2E-02	y
rs7359257	IQCH	intronic	15	65489961	0.45	A	C	C	1.73	0.36	1.9E-06	C	2.9E-02	y
rs17268785	CCDC85A	intronic	2	56445587	0.17	G	A	A	3.22	0.50	9.7E-11	G	9.3E-02	n
rs9635759	CA10	~94kb	17	46968784	0.32	A	G	G	2.97	0.41	7.3E-13	G	1.0E-01	y
rs9555810	C13orf16, ARHGEF7	~185kb, ~223kb	13	110979438	0.28	G	C	C	2.31	0.43	5.6E-08	C	1.1E-01	y
rs1862471	OLFM2	intronic	19	9861322	0.47	G	C	C	2.03	0.39	1.5E-07	C	1.3E-01	y
rs2090409	TMEM38B	~400kb	9	108006909	0.31	A	C	A	4.73	0.39	2.2E-33	A	1.6E-01	y
rs757647	KDM3B	intronic	5	137735214	0.22	A	G	A	2.40	0.44	5.4E-08	A	1.7E-01	y
rs1398217	FUSSEL18	intronic	18	43006236	0.43	G	C	G	2.69	0.37	2.3E-13	C	2.2E-01	n
rs6439371	TMEM108, NPHP3	~146kb, ~170kb	3	134093442	0.34	G	A	A	2.32	0.41	1.3E-08	G	2.3E-01	n
rs12617311	PLCL1	~195kb	2	199340810	0.32	A	G	A	3.03	0.42	6.0E-13	G	2.3E-01	n
rs10980926	ZNF483	intronic	9	113333455	0.36	A	G	G	2.49	0.38	4.2E-11	A	2.4E-01	n
rs7821178	PXMP3	~181kb	8	78256392	0.34	A	C	A	2.36	0.40	3.0E-09	C	2.8E-01	n
rs1361108	C6orf173, TRMT11	~98kb, ~407kb	6	126809293	0.46	T	C	T	2.13	0.38	1.7E-08	C	3.0E-01	n
rs2243803	SLC14A2	~238kb	18	41210670	0.40	A	T	T	1.96	0.38	3.4E-07	T	3.3E-01	y
rs900145	ARNTL	~5 kb	11	13250481	0.30	C	T	T	2.34	0.41	1.6E-08	T	3.5E-01	y
rs1659127	MKL2	~28kb	16	14295806	0.34	A	G	G	2.43	0.41	4.0E-09	A	3.5E-01	n
rs12472911	LRP1B	intronic	2	141944979	0.20	C	T	T	2.54	0.48	1.5E-07	T	3.6E-01	y
rs1079866	INHBA	~250kb	7	41436618	0.15	G	C	C	3.95	0.52	5.5E-14	G	4.4E-01	n
rs7759938	LIN28B	~26kb	6	105485647	0.32	C	T	T	6.45	0.39	5.4E-60	C	4.4E-01	n
rs2002675	TRA2B, ETV5	~4kb, ~135kb	3	187112262	0.42	G	A	A	2.23	0.37	1.2E-09	G	4.6E-01	n
rs10423674	CRTC1	intronic	19	18678903	0.35	A	C	C	2.28	0.39	5.9E-09	C	5.0E-01	y
rs6589964	BSX	~18kb	11	122375893	0.48	A	C	A	2.68	0.38	1.9E-12	A	5.3E-01	y
rs2687729	EEFSEC	intronic	3	129377916	0.27	G	A	A	2.29	0.43	1.3E-07	A	5.3E-01	y
rs1364063	NFAT5	~10kb	16	68146073	0.43	C	T	T	2.13	0.38	1.8E-08	T	5.6E-01	y
rs17188434	NR4A2	~84kb	2	156805022	0.07	C	T	C	4.52	0.74	1.1E-09	C	5.9E-01	y
rs7642134	VGLL3	~70kb	3	86999572	0.38	A	G	A	2.41	0.38	3.5E-10	G	7.2E-01	n
rs6438424	3q13.32	intergenic	3	119057512	0.50	A	C	A	2.73	0.37	1.4E-13	A	7.3E-01	y
rs4840086	PRDM13, MCHR2	~145kb, ~160kb	6	100315159	0.42	G	A	G	2.11	0.38	2.4E-08	G	7.6E-01	y
rs7617480	KLHDC8B	intronic	3	49185736	0.22	A	C	C	2.42	0.44	2.8E-08	A	8.4E-01	n
rs6575793	BEGAIN	intronic	14	100101970	0.42	C	T	T	2.27	0.40	1.2E-08	T	8.6E-01	y
rs852069	PCSK2	~84kb	20	17070593	0.37	A	G	A	2.08	0.38	3.3E-08	G	8.8E-01	n
rs3743266	RORA	3'UTR	15	58568805	0.32	C	T	C	2.01	0.41	8.0E-07	T	8.9E-01	n
rs16938437	PHF21A	intronic	11	46009151	0.09	T	C	T	3.67	0.68	5.9E-08	C	9.7E-01	n
rs6762477	RBM6	intronic	3	50068213	0.44	G	A	A	2.54	0.45	1.6E-08	*data for rs6762477 unavailable in GIANT*		

[a] Indicates whether menarche decreasing and BMI increasing alleles are consistent (y=yes, n=no)

Supplementary Table 16: Association of menarche loci with height in ~130,000 adults in the GIANT consortium.

								Association with menarche				Association with adult height				
SNP	Nearest gene(s)	Distance from gene (kb)	Chr	Position (B36)	MAF	Minor allele	Major allele	Menarche decreasing allele	Beta (weeks/allele)	se	P-value	Height decreasing allele	Zscore per allele	se	P-value	Direction consistent[a]
rs7759938	LIN28B	~26kb	6	105485647	0.32	C	T	T	6.45	0.39	5.4E-60	T	0.042	0.005	8.7E-18	y
rs1361108	C6orf173, TRMT11	~98kb, ~407kb	6	126809293	0.46	T	C	T	2.13	0.38	1.7E-08	C	0.035	0.005	1.9E-14	n
rs7821178	PXMP3	~181kb	8	78256392	0.34	A	C	A	2.36	0.40	3.0E-09	C	0.023	0.005	1.6E-06	n
rs1659127	MKL2	~28kb	16	14295806	0.34	A	G	G	2.43	0.41	4.0E-09	G	0.024	0.005	2.9E-06	y
rs2090409	TMEM38B	~400kb	9	108006909	0.31	A	C	A	4.73	0.39	2.2E-33	A	0.020	0.005	2.8E-05	y
rs1364063	NFAT5	~10kb	16	68146073	0.43	C	T	T	2.13	0.38	1.8E-08	T	0.018	0.005	5.9E-05	y
rs2002675	TRA2B, ETV5	~4kb, ~135kb	3	187112262	0.42	G	A	A	2.23	0.37	1.2E-09	A	0.017	0.005	2.2E-04	y
rs16938437	PHF21A	intronic	11	46009151	0.09	T	C	T	3.67	0.68	5.9E-08	T	0.027	0.008	9.5E-04	y
rs10980926	ZNF483	intronic	9	113333455	0.36	A	G	G	2.49	0.38	4.2E-11	G	0.015	0.005	1.7E-03	y
rs6589964	BSX	~18kb	11	122375893	0.48	A	C	A	2.68	0.38	1.9E-12	A	0.014	0.005	2.5E-03	y
rs6438424	3q13.32	intergenic	3	119057512	0.50	A	C	A	2.73	0.37	1.4E-13	A	0.013	0.004	4.0E-03	y
rs900145	ARNTL	~5 kb	11	13250481	0.30	C	T	T	2.34	0.41	1.6E-08	T	0.013	0.005	9.4E-03	y
rs1079866	INHBA	~250kb	7	41436618	0.15	G	C	C	3.95	0.52	5.5E-14	C	0.016	0.006	1.5E-02	y
rs17188434	NR4A2	~84kb	2	156805022	0.07	C	T	C	4.52	0.74	1.1E-09	C	0.021	0.009	1.7E-02	y
rs3914188	ECE2	3'UTR	3	185492742	0.27	G	C	G	2.21	0.43	2.6E-07	C	0.012	0.005	2.8E-02	n
rs4840086	PRDM13, MCHR2	~145kb, ~160kb	6	100315159	0.42	G	A	G	2.11	0.38	2.4E-08	G	0.010	0.005	3.2E-02	y
rs3743266	RORA	3'UTR	15	58568805	0.32	C	T	C	2.01	0.41	8.0E-07	C	0.010	0.005	4.2E-02	y
rs13187289	PHF15	~12kb	5	133877076	0.20	G	C	C	3.02	0.48	1.9E-10	C	0.012	0.006	4.3E-02	y
rs1398217	FUSSEL18	intronic	18	43006236	0.43	G	C	G	2.69	0.37	2.3E-13	G	0.009	0.005	6.1E-02	y
rs9939609	FTO	intronic	16	52378028	0.40	A	T	A	2.12	0.39	3.1E-08	A	0.008	0.005	6.4E-02	y
rs12472911	LRP1B	intronic	2	141944979	0.20	C	T	T	2.54	0.48	1.5E-07	C	0.010	0.006	6.6E-02	n
rs12617311	PLCL1	~195kb	2	199340810	0.32	A	G	A	3.03	0.42	6.0E-13	A	0.007	0.005	2.0E-01	y
rs633715	SEC16B	~44kb	1	176119203	0.20	C	T	C	2.63	0.47	2.1E-08	T	0.007	0.006	2.2E-01	n
rs2243803	SLC14A2	~238kb	18	41210670	0.40	A	T	T	1.96	0.38	3.4E-07	T	0.005	0.005	2.7E-01	y
rs2947411	TMEM18	~53kb	2	604168	0.17	A	G	G	2.84	0.51	1.7E-08	A	0.006	0.006	3.0E-01	n
rs17268785	CCDC85A	intronic	2	56445587	0.17	G	A	A	3.22	0.50	9.7E-11	A	0.006	0.006	3.1E-01	y
rs4929923	TRIM66	3'UTR	11	8595776	0.36	T	C	C	2.27	0.40	1.2E-08	C	0.004	0.005	3.6E-01	y
rs9635759	CA10	~94kb	17	46968784	0.32	A	G	G	2.97	0.41	7.3E-13	G	0.005	0.006	4.1E-01	y
rs466639	RXRG	intronic	1	163661506	0.13	T	C	T	4.20	0.57	1.3E-13	T	0.005	0.007	4.8E-01	y
rs10423674	CRTC1	intronic	19	18678903	0.35	A	C	C	2.28	0.39	5.9E-09	C	0.003	0.005	4.9E-01	y
rs6439371	TMEM108, NPHP3	~146kb, ~170kb	3	134093442	0.34	G	A	A	2.32	0.41	1.3E-08	G	0.003	0.005	5.4E-01	n
rs1862471	OLFM2	intronic	19	9861322	0.47	G	C	C	2.03	0.39	1.5E-07	C	0.003	0.005	5.7E-01	y
rs852069	PCSK2	~84kb	20	17070593	0.37	A	G	A	2.08	0.38	3.3E-08	G	0.002	0.005	6.2E-01	n
rs7359257	IQCH	intronic	15	65489961	0.45	A	C	C	1.73	0.36	1.9E-06	A	0.002	0.005	6.3E-01	n
rs757647	KDM3B	intronic	5	137735214	0.22	A	G	A	2.40	0.44	5.4E-08	G	0.003	0.005	6.3E-01	n
rs10899489	GAB2	intronic	11	77773021	0.15	A	C	C	3.09	0.54	8.1E-09	C	0.003	0.006	6.5E-01	y
rs7642134	VGLL3	~70kb	3	86999572	0.38	A	G	A	2.41	0.38	3.5E-10	A	0.002	0.005	6.8E-01	y
rs7617480	KLHDC8B	intronic	3	49185736	0.22	A	C	C	2.42	0.44	2.8E-08	A	0.002	0.005	7.4E-01	n
rs6575793	BEGAIN	intronic	14	100101970	0.42	C	T	T	2.27	0.40	1.2E-08	T	0.001	0.005	8.2E-01	y
rs9555810	C13orf16, ARHGEF7	~185kb, ~223kb	13	110979438	0.28	G	C	C	2.31	0.43	5.6E-08	G	0.001	0.005	8.5E-01	n
rs6762477	RBM6	intronic	3	50068213	0.44	G	A	A	2.54	0.45	1.6E-08	G	0.001	0.006	8.6E-01	n
rs2687729	EEFSEC	intronic	3	129377916	0.27	G	A	A	2.29	0.43	1.3E-07	G	0.001	0.005	9.0E-01	n

[a] Indicates whether menarche decreasing and height decreasing alleles are consistent (y=yes, n=no)

Supplementary Table 17: Association of menarche loci with height in ~130,000 adults in the GIANT consortium.

	n	beta (weeks/allele)	SE	p
Menarche risk allele score	5721	-1.76	0.30	4.6E-09
adjusted for mother's age	5721	-1.76	0.30	4.1E-09
adjusted for mother's age and BMI	5381	-1.74	0.30	9.4E-09
adjusted for mother's age and height	5641	-1.69	0.30	1.7E-08
adjusted for mother's age, BMI and height	5381	-1.73	0.30	9.0E-09

*Calculated as the sum of menarche-lowering alleles in each individual across 30 SNPs
Missing genotypes were imputed with the mean value in individuals with at least 18 succesfully genotyped SNPs

SUPPLEMENTARY NOTE

CONTENTS

STUDY INFORMATION

a) Stage 1 GWAS studies

AGES-Reykjavik Study: The Reykjavik Study cohort originally comprised a random sample of 30,795 men and women born in 1907-1935 and living in Reykjavik in 1967. A total of 19,381 people participated in the Reykjavik Study examination, a 71% recruitment rate. The study sample was divided into six groups by birth year and birth date within month. One group was invited to participate in all subsequent examinations, while one group was designated as a control group and was not included in examinations until 1991. Other groups were invited to participate in specific examinations of the study. Between 2002 and 2006, the AGES-Reykjavik Study re-examined 5764 survivors of the original Reykjavik Study. Successful genotyping was available for 1849 AGES women participants who were eligible for this study. The AGES-Reykjavik Study GWAS was approved by the National Bioethics Committee and the Data Protection Authority and also was covered under the MedStar Institutional Review Board. All subjects provided written informed consent.

Amish: The Old Order Amish are a closed founder population in Lancaster County, PA. Due to their unique immigration and ancestral history, all current living Old Order Amish can be connect via one single 13 generation pedigree[1,2]. Women included in this study were recruited as part of at least one of multiple studies performed in this population and are described elsewhere[3-7]. All study protocols were approved by the Institutional Review Board at the University of Maryland, Baltimore and informed consent was obtained from each participant. Physical examinations were performed at the Amish Research Clinic in Strasburg, PA and a reproductive health questionnaire was completed by self report. Presenting pregnant or within 6 months postpartum were common exclusion criteria among all study designs.

ARIC: The ARIC study is a multi-center prospective investigation of atherosclerotic disease in a bi-racial population[8]. Men and women aged 45-64 years at baseline were recruited from 4 communities: Forsyth County, North Carolina; Jackson, Mississippi; suburban areas of Minneapolis, Minnesota; and Washington County, Maryland. A total of 15,792 individuals participated in the baseline examination in 1987-1989, with four follow-up examinations in approximate 3-year intervals, during 1990-1992, 1993-1995, and 1996-1998. Only White women with genotype data and age at menarche between 9 and 17 years of age were included in this analysis (N=4247). This study was approved by the institutional review board at each field center, and this analysis was approved by the University of North Carolina at Chapel Hill School of Public Health Institutional Review Board on research involving human subjects. All subjects provided written informed consent.

The 1958 British Birth Cohort (B58C): The 1958 Birth Cohort (also known as the National Child Development Study) is a national population sample followed periodically from birth to age 44-45 years. It includes all births in England, Wales and Scotland, during one week in 1958. Age at menarche was derived from reports at examination at the age of 16[9]. Genotyping of this study was previously performed as a part of the Wellcome Trust Case Control Consortium (WTCCC) and the Type 1 Diabetes Genetics Consortium (T1DGC) and has been described previously[10-12]. The current analysis included 1584 individuals that passed the quality control criteria and that had data on age at menarche.

Cohorte LAUSannoise (CoLaus): CoLaus is a cross-sectional study of a random sample of 6188 European adults (including 2,976 women), aged 35–75 years, living in Lausanne, Switzerland[13]. Recruitment took place between April 2003 and March 2006. Only individuals with four grandparents of European origin were included in the study. Participants provided a detailed health questionnaire, underwent a physical exam and donated blood after a 12-hr fasting period for clinical chemistry and genetic analyses. Following exclusions due to quality control criteria and missing data, 2,874 women were included in the genome-wide analyses. In all studies, age at menarche to the nearest completed whole year was ascertained at baseline by questionnaire. The CoLaus study was sponsored in part by GlaxoSmithKline, and all participants were duly informed about this sponsorship, and consented for the use of biological samples and data by GlaxoSmithKline and its subsidiaries; the study was approved by the Local Ethics committee.

deCODE Genetics (Iceland): Self-reported age at menarche was available from 39,728 Icelanders. The information had been collected in a nationwide cancer screening program through the Cancer Detection Clinic at the Icelandic Cancer society since 1964. The question referred to age at previous birthday before onset of the first menstruations and the individuals had a reported age at first menstruation between 8 and 20 years. Of these individuals, 15,864 were genotyped on an Illumina 317K/ 370 K SNP chip in one of several genome-wide association studies recently conducted by deCODE Genetics. All of these studies were approved by the Data Protection Commission of Iceland and the National Bioethics Committee of Iceland. Written informed consent was obtained from all participants. Personal identifiers associated with phenotypic information and blood samples were encrypted using a third-party encryption system as previously described. Only individuals with a genotype yield over 98% were included in the study.

Danish National Birth Cohort: DNBC is a population-based cohort of 101,042 pregnancies, recruited in the years 1996-2002[14]. All participating women underwent thorough phenotype characterization based on information from four computer-assisted telephone interviews conducted during pregnancy (two interviews) and after delivery (two interviews). During the first interview women were asked “How old were you when you had your first menstrual period?” GWAS data were generated for 3,840 individuals from the DNBC (mothers

and their children) in a study of prematurity and its complications (Principal investigator Jeff Murray) within the Gene Environment Association Studies (GENEVA) consortium. Age at menarche between 9 and 17 years and genome-wide genotype and imputed data were available for 1,748 women. The DNBC study protocol was approved by the Danish Scientific Ethical Committee and the Danish Data Protection Agency.

Estonian Genome Center, University of Tartu (EGCUT): The Estonian cohort is from the population-based biobank of the Estonian Genome Project of University of Tartu. The project was conducted according to the Estonian Gene Research Act and all participants signed the broad informed consent[15] (www.geenivaramu.ee). The current cohort size is over 43,000, from 18 years of age and up, which reflects closely the age distribution in the adult Estonian population. Participants were randomly selected from individuals visiting GP offices or hospitals and were recruited by general practitioners (GP) and physicians. Each participant filled out a Computer Assisted Personal interview, which included personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, three generations), educational and occupational history and lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, women´s health, quality of life). Anthropometric and physiological measurements were also taken. GWAS was performed on 2,700 randomly selected individuals[16] with the Illumina HumanHapCNV370 array, according to the Illumina protocol (www.illumina.com) in the Estonian Biocenter Genotyping Core Facility. Age at menarche between 9 and 17 years were available for 987 females for the genome wide meta-analyis stage of this project. Data on a further 196 women was available for in silico replication.

EPIC-Norfolk Obesity Case-Cohort: The European Prospective Investigation into Cancer and Nutrition (EPIC-Norfolk) cohort was initially designed to investigate the relationship between diet, cancer and chronic disease. The total study size was 25,639 men and women of European descent from Norfolk in the United Kingdom and aged between 39 and 79 years in 1993-1997[17]. The EPIC Obesity case-cohort study includes obesity cases and cohort controls. 1685 obese cases (718 women), defined as those with a body mass index >30 kg/m^2 , were randomly selected from the obese individuals within EPIC-Norfolk. The control-cohort consists of a further 2566 individuals (1364 women) randomly selected from the EPIC-Norfolk study. Age at menarche in completed whole years was ascertained at baseline in women by questionnaire. Following exclusions due to quality control criteria and missing data, data on 625 obese and 1,215 control women were available for genome-wide analysis. Ethical approval for the study was granted by the Norwich Local Research Ethics Committee. All subjects gave written informed consent.

ERF: The Erasmus Rucphen Family study is part of the Genetic Research in Isolated Population program. The study population essentially consists of one extended family of descendents from 20 related couples who lived in the isolate between 1850 and 1900 and had at least 6 children baptized in the community church. The

detailed information about ERF isolate can be found elsewhere[18]. The Medical Ethical Committee of the Erasmus Medical Center, Rotterdam approved the study and informed consent was obtained from all participants. Self-reported age at start of menarche was assessed by a questionnaire and genome-wide and imputed data were available for 1103 women.

The Framingham Heart Study began in 1948 to study determinants of cardiovascular disease and other major medical conditions[19,20]. In 1971, Offspring of the Original Cohort participants and Offspring spouses were enrolled into the Framingham Offspring Study. Offspring participants, including 2641 women (mean age 36 years), have been examined approximately every 4 to 8 years[21,22]. From 2002 to 2005, 4095 adults including 2182 women (mean age 40 years) with at least one parent in the Offspring Study were enrolled in the Framingham Third Generation cohort[23]. Women were queried about menarche at the second Offspring examination (1979 to 1982) and at the first Third Generation examination with the following questions: "*Age at start of menses*" and "*How old were you when you had your first menstrual period (menses)?"*, respectively. The self-reported age at first period was recorded. Offspring women participating in the Framingham Osteoporosis Study (1996 – 2001) were asked about menarche with the following query: *"About how old were you when you had your first menstrual period?"*. The self-reported data from the Osteoporosis examination was used (n=214) if menarche data was not available from Offspring examination two. There were 1777 Offspring Cohort and 2024 Third Generation women who reported an age at menarche between 9 and 17 years with genotyping available.

SNP weights for 10 principal components (PCs) were inferred using a maximal set of independent individuals; the PCs for the remaining individuals were computed using the SNP weights obtained from the unrelated set of individuals. The first PC (PC1) was significantly associated with age at menarche ($P<0.01$), and therefore was included as a covariate in all SNP association analyses. In addition, we adjusted for birth cohort by decade. Linear mixed effects models were used to account for familial correlations. Each SNP was tested for association with age at menarche using an additive genetic model.

HBCS: Helsinki Birth Cohort Study (HBCS) includes 8760 subjects born in Helsinki between 1934 and 1944. A representative subset of 928 males and 1075 females participated in a clinical study between 2000 and 2002 focusing on cardiovascular and metabolic outcomes. During the clinical visit information on age at menarche was collected with a questionnaire[24]. The current study is restricted to 976 females for which successful genotyping on Illumina 670 Quad arrays (modified from Illumina Infinium 610K arrays) and data on age of menarche were available.

Health 2000 cases and controls: The Health 2000 (H2000), is a health interview/examination survey carried out by the National Institute for Health and Welfare in Finland from fall 2000 to spring 2001, with a nationally representative sample of 10,000 individuals drawn from the Finnish population aged 18 and older. The main topics of the study were health status, major chronic conditions, functional ability and limitations, determinants of health, and use of health care[25]. Data on female reproductive health, including age at menarche, was collected with a separate questionnaire. From a sub-cohort of 6,000 individuals representative of the Finnish population over age 30, roughly 1000 non-diabetic subjects meeting the IDF criteria for metabolic syndrome and control cohort of 1000 subjects matched for sex, age and residence, were selected for GWAS analyses. The current study is restricted to 457 females from the control group and 465 females from the metabolic syndrome case group for which successful genotyping on Illumina 670 Quad arrays (modified from Illumina Infinium 610K arrays) and data on age of menarche were available.

For all study subjects participating in the different Finnish cohort studies (HBCS, NFBC and Health 2000), informed consent was obtained using protocols approved by the corresponding local Ethical Committees. Furthermore, the studies were approved by the national Data Protection Boards where relevant.

InCHIANTI: The InCHIANTI study is a population-based epidemiological study aimed at evaluating factors that influence mobility in the older population living in the Chianti region of Tuscany, Italy. Details of the study have been previously reported[26]. Briefly, 1616 residents were selected from the population registry of Greve in Chianti (a rural area; 11,709 residents with 19.3% of the population greater than 65 years of age) and Bagno a Ripoli (Antella village near Florence; 4704 inhabitants, with 20.3% greater than 65 years of age). The participation rate was 90% (n= 1453) and participants ranged between 21–102 years of age. The study protocol was approved by the Italian National Institute of Research and Care of Aging Institutional Review. There were 85 parent-offspring pairs, 6 sib-pairs and 2 half-sibling pairs documented. We investigated any further familial relationships using IBD of 10,000 random SNPs using RELPAIR and uncovered 1 parent-offspring, 79 siblings and 13 half-sibling[27]. We utilized the correct family structure inferred from genetic data for all analyses.

Indiana: Genetic studies of bone density and related phenotypes in premenopausal women at Indiana have been ongoing since 1988, beginning with studies in twin pairs. This has expanded over time to include sibling pair linkage and association studies[28]. Peak bone mineral density as measured premenopausally in women and before age 60 in men is the primary quantitative phenotype of interest, specifically at the femoral neck and lumbar spine. The sample consists of European-American premenopausal sister pairs from Indiana, at least 20 years of age. The subjects were recruited without regard to bone density or other clinical phenotype. The exclusion criteria were limited to irregular menses or a history of pregnancy or lactation within three months

prior to enrollment, a history of chronic disease, current medications known to affect bone mass or metabolism, or inability to have BMD measured due to obesity. GWAS genotyping was performed by CIDR using the Illumina HumanHap 610 Quad version 1B platform, and the current analysis included all individuals with GWAS genotypes and age of menarche as measured by recall at their study visit.

Details of the **Nurses' Health Study (NHS)** cohorts have been described previously[29]. Briefly, the NHS was initiated in 1976, when 121,700 United States registered nurses between the ages of 30 and 55, residing in 11 larger U.S. states, returned an initial questionnaire reporting medical histories and baseline health-related exposures, including information related to reproductive history (age at menarche, age at first birth, parity, age at menopause etc.), and exposure to exogenous hormones (oral contraception or post-menopausal hormone replacement therapy). Biennial questionnaires with collection of exposure information on risk factors have been collected prospectively, and outcome data with follow-up of reported disease events are collected. From May 1989 through September 1990, we collected blood samples from 32,826 participants in the NHS cohort. Subsequent follow-up has been greater than 99% for this subcohort. Informed consent was obtained from all participants. The study was approved by the Institutional Review Board of the Brigham and Women's Hospital, Boston, MA, USA.

NHS breast cancer GWAS (CGEMS): The NHS nested breast cancer case-control study was derived from the 32,826 women in the blood subcohort who were free of diagnosed breast cancer at blood collection and followed for incidence disease until June 1, 2004. Breast cancer follow-up in the NHS was conducted by personal mailings and searches of the National Death Index. Controls were women not diagnosed with breast cancer during follow-up, and were one-to-one matched to cases based on age at diagnosis, blood collection variables (time of day, season, and year of blood collection, as well as recent (<3 months) use of postmenopausal hormones), ethnicity (all cases and controls are self-reported Caucasians), and menopausal status (all cases were postmenopausal at diagnosis). The 2,287 NHS participants included in the present analysis were from this nested breast cancer case-control study and were self-described Caucasians with genotype data available from the National Cancer Institute's Cancer Genetic Marker of Susceptibility (CGEMS) project[30]. There were 2270 women reporting age at menarche between ages 9-17 years with genotyping data available.

NHS type 2 diabetes (T2D) GWAS: NHS participants for the current T2D GWAS were also selected among those with a blood sample using a nested case-control design[31]. Diabetes cases were defined as self-reported diabetes confirmed by a validated supplementary questionnaire. For cases before 1998, diagnosis was made using criteria consistent with those proposed by the National Diabetes Data Group (NDDG)[32]. We used the American Diabetes Association diagnostic criteria for diagnosis of diabetes cases during the 1998 and 2000

cycles[33]. 98% of self-reported cases were confirmed by medical records review in this cohort[34]. Controls were defined as those free of diabetes at the time of diagnosis of the case and remained unaffected through follow-up (2006). Although controls were originally matched per case (by gender, year of birth, month of blood collection, and fasting status), matched pairs were broken because not all subjects gave informed consent for submission of their GWAS data to dbGaP. The current analysis included 3090 Caucasian women who had genotyping data and reported age at menarche between ages 9-17 years.

The NHS breast cancer GWA scan used the Illumina Infinium Sentrix HumanHap550 chip. Detailed methods related to the genotyping have been published previously[30]. Imputation of untyped genotypes was based on HapMap haplotype reference (release 21) using MACH software.

The NHS T2D GWA scan is a component of the Gene Environment-Association Studies (GENEVA) under the NIH Genes, Environment and Health Initiative (GEI). Genotyping was done at the Broad Center for Genotyping and Analysis using the Affymetrix Human 6.0 array (Santa Clara, CA) and the Birdseed calling algorithm[35]. Imputation of untyped genotypes was based on HapMap haplotype reference (release 22) using MACH software.

NFBC: The Northern Finland Birth Cohort 1966 (NFBC1966; http://kelo.oulu.fi/NFBC/pub/) is a prospective cohort study conducted in the two northernmost provinces in Finland, Oulu and Lapland. The study enrolled mothers living in the district who had estimated dates of delivery in year 1966. Altogether 12,231 births were included in the study, representing 96% of all births in these provinces. Data on the mothers' height and pre-pregnancy weight were collected from standard forms for the pregnancy follow-up and the maternity cards carried by all mothers. The offspring were followed up at 6 months, 1, 14 and 31 years of age. Data on age at menarche was obtained retrospectively by a postal questionnaire at age 31. Also at age 31, a representative sub-sample of the study subjects (cohort members still living in Northern Finland or in the capital area) were invited for clinical examination. At the clinical visit, height and weight were assessed. The participants (n=5654), also provided a fasting blood sample, from which DNA was extracted. The study protocol, the data collection protocol and the clinical measurement procedures of the complete cohort study have earlier been described in detail[36,37]. The current study is restricted to females for which successful genotyping was completed on Illumina Infinium 370cnvDuo arrays, as described by Sabatti et al (2009)[38], and for which age at menarche data were available.

NTR: As part of a longitudinal survey study of health, lifestyle and personality, twins and their family members registered with the Netherlands Twin Register (NTR) are approached every 2 to 3 years[39]. As part of a case-control study for major depression disorder, genotype information was obtained in 1940 NTR participants[40]. In

the surveys female participants were asked at what age they had their first menstrual period. Inconsistencies over time were checked. Age of menarche was available for 1051 participants.

QIMR: Data for age at menarche were available from two separate cohorts referred to as the Adult and Adolescent cohorts. The adult data were gathered from a number of studies carried out between 1980 and 1995. Many of these studies were follow-up studies and so there is a large overlap in participants. There are two main Adult cohorts. The first study, known as The Canberra Study, involved twins from the Australian Twin Registry who were aged between 17 and 88 when they were surveyed between 1980 and 1982. The twins were mailed a health questionnaire which included a number of questions about their reproductive history including a question about their age at menarche -"How old were you (in years and months) when you had your first menstrual period? A second cohort was recruited in 1989 involving twins born between 1964 and 1971 and their first degree relatives. They answered a similar questionnaire to the one answered by the Canberra cohort. The adult cohorts provided a total of 2,256 individuals with age at menarche to the analysis

Adolescent twins and their siblings were drawn from the adolescent cohorts that had been recruited as part of ongoing studies of melanoma risk factors[41] and cognition[42]. The protocol for obtaining information on age at menarche has been described in a previous paper[43]. Due to additional data collection, this study includes more individuals than the previous linkage analysis. The young age of the cohort meant that there was a significant amount of censoring in the data. All censored individuals who were younger than the mean age at menarche in the sample at the age of their last interview, were removed from the analysis. After removal of non-genotyped individuals, individuals who were removed in the genotypic QC process, and those who were phenotypic outliers there were 1,274 individuals with phenotype and genotype information remaining. Of these, 2 individuals had a censored age at menarche and these were excluded from the analysis. In cases where there was more than one measure of age at menarche, the earliest measure was used as this was assumed to be closest to the true age, and hence the most accurate. Details of the genotyping and imputation are given elsewhere[44].

Rotterdam Study I, II, and III are ongoing prospective population-based cohort studies, focus on chronic disabling conditions of the elderly in the Netherlands. In summary, men and women aged 55 years or older, living in Ommoord, a suburb of Rotterdam, the Netherlands, were invited to participate[45]. Self-reported age at menarche was assessed by questionnaire. Age at menarche between 9 and 17 years (collected retrospectively) and genome-wide genotype and imputed data were available for 3175 (RSI), 1119 (RSII) and 1112 (RS3) women. GWA analysis was performed using GRIMP[46].

SAGE: The Study of Addiction: Genetics and Environment is funded as part of the Gene Environment Association Studies initiative (GENEVA) supported by the National Human Genome Research Institute (dbGaP

study accession phs000092.v1.p1). Subjects were selected from three large, complementary datasets: the Collaborative Study on the Genetics of Alcoholism (COGA), the Family Study of Cocaine Dependence (FSCD), and the Collaborative Genetic Study of Nicotine Dependence (COGEND). The Institutional Review Board at each contributing institution reviewed and approved the protocols for genetic studies under which all subjects were recruited. All subjects completed a comprehensive psychiatric interview that was based on the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA)[47]. As part of this assessment, female subjects were asked "At what age did you have your first menstrual period?" Data are available for 1376 subjects; mean age at menarche was 12.8. For additional description of the studies, see http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000092.v1.p1.

Samples were genotyped at the Johns Hopkins Center for Inherited Disease Research (CIDR). Data were released for 4,189 study samples. Study samples, including 49 study duplicates, were plated and genotyped together with 135 HapMap controls (86 CEU; 49 YRI). Genotyping was performed using Illumina Human1Mv1_C BeadChips (Illumina, San Diego, CA, USA) and the Illumina Infinium II assay protocol[48]. Allele cluster definitions for each SNP were determined using Illumina BeadStudio Genotyping Module version 3.1.14 and the combined intensity data from the samples. Strict quality control standards were implemented and genotypes were released by CIDR for 1,040,106 SNPs (99.15% of attempted). The mean non-Y SNP call rate and mean sample call rate was 99.7% for the released CIDR dataset. Study duplicate reproducibility was 99.98%. Further extensive cleaning was undertaken to insure high quality genotyping by examining batch effects, potential chromosomal anomalies, and Mendelian errors (Laurie et al., under review).

SNPs with an allele frequency > 1% in either the European or African descent populations were analyzed (948,658 SNPs). A SNP call rate of 98% was required. Hardy Weinberg equilibrium (HWE) was tested and SNPs that deviated from HWE ($p < 10–4$) were excluded. The final number of subjects included in analyses was 3,834. Individuals were dropped if there was potential sample misidentification, sample relatedness, or other misspecification (N=171).

SardiNIA: The SardiNIA genome wide association study has been described in detail previously[49,50]. Briefly, the GWA study examined a total of 4,305 related individuals participating in a longitudinal study of aging-related quantitative traits in the Ogliastra region of Sardinia, Italy. Genotyped individuals had four Sardinian grandparents and were selected for genotyping without regard to their phenotypes. Among the individuals examined, 1,412 were genotyped with the Affymetrix Mapping 500K Array Set. A total of 356,359 autosomal SNPs met the quality control criteria and were used as input for the imputation procedure using the software MACH[51,52]. The remaining 2,893 individuals were genotyped with the Affymetrix Mapping 10K Array. These individuals were mostly offspring and siblings of the 1,412 individuals that were genotyped with the Affymetrix

Mapping 500K Array Set. We took advantage of the relatedness among individuals to impute missing genotypes in these additional individuals; we identified large stretches of chromosome shared within each family and probabilistically "filled-in" genotypes within each stretch whenever one or more of its carriers was genotyped with the 500K Array Set[51,53]. In order to more efficiently evaluate identity-by-descendent states at non-overlapping markers, 436 individuals out of the 1,412 were also genotyped with the 10K Array. Among the 4,305 genotyped individuals, 2158 women were analysed for age at menarche.

TwinsUK: The TwinsUK cohort consisted of a group of twins ascertained to study the heritability and genetics of age-related diseases (www.twinsUK.ac.uk). These unselected twins were recruited from the general population through national media campaigns in the UK and shown to be comparable to age-matched population singletons in terms of disease-related and lifestyle characteristics[54,55].
The TwinsUK II and III cohorts consist of twins from the adult twin British registry, also shown to be representative of singleton populations and the United Kingdom population[56]. Age at menarche was assessed by questionnaire and genome-wide genotype and imputed data were available for 2276 (TwinsUK), 671 (TwinsUKII) and 1016 (TwinsUKIII) women. Ethics approval was obtained from the Guy's and St. Thomas' Hospital Ethics Committee. Written informed consent was obtained from every participant to the study.

TwinsUK samples were typed with the Infinium 610k assay (Illumina, San Diego, USA) at two different centers, and namely the Center for Inherited Diseases Research (USA) and the Wellcome Trust Sanger Institute. We pooled the normalised intensity data and called genotypes on the basis of the Illluminus algorithm. No calls were assigned if the most likely call was less than a posterior probability of 0.95. Validation of pooling was done by visual inspection of 100 random, shared SNPs for overt batch effects; none were observed. We excluded SNPs that had a low call rate (≤ 90%), Hardy-Weinberg p values < 10−4 and minor allele frequencies < 1%. We, also removed subjects where genotyping failed for >2 % of SNPs. The overall genotyping efficiency of the GWA was 98.7 %. Imputation of genotypes was carried out using the software IMPUTE[57].

WGHS: The Women's Genome Health Study (WGHS) is a prospective cohort of female healthcare professionals, aged 45 or older at baseline, who provided baseline blood sample and consent for blood based analysis in the Women's Health Study (WHS), a randomized, placebo controlled trial of aspirin and vitamin E in the primary prevention of cardiovascular disease and cancer. A complete description of the WGHS has been published previously[58].

All information about reproductive aging was determined by self-report questionnaire at baseline and subsequent follow-up . For age of menarche, participants were asked "''At what age did your menstrual periods begin?'' with response categories ''9 or younger; 10; 11; 12; 13; 14; 15; 16; 17 or older.'' For age of

natural menopause, participants were asked ''Have your menstrual periods ceased permanently?'' If yes, ''At what age did your natural periods cease?'' and ''For what reason did your periods cease?'' with response categories ''Surgical; Radiation or Chemotherapy; Natural.'' Age at natural menopause was assessed in the baseline questionnaire for postmenopausal women at baseline, and updated in subsequent questionnaires for premenopausal women at baseline. In the current analysis all WGHS participants had passed through menopause. Women who reported menopause before 40 or after 60 were excluded.
Association testing for reproductive aging phenotypes was performed with Mach2Qtl v. 1.0.4.

Among WGHS participants, genotyped data was collected using the Illumina HumanHap300 Duo "+" platform, for which the "+" or custom content included SNPs chosen for suspected biological consequences as well as for increased coverage of genes related to cardiovascular disease. In final genotype data included 22,054 participants whose self-reported European ancestry was confirmed by principal component analysis in PLINK[59]. All of these samples had genotype information for >98% of the SNPs; all SNPs had complete information for >90% of the samples and deviations from Hardy-Weinberg equilibrium not exceeding $p < 10^{-6}$ in significance. Imputation of genotypes for a total of > 2.5 million SNPs was performed with MACH v. 1.0.16 (http://www.sph.umich.edu/csg/abecasis/MaCH/index.html) using linkage disequilibrium relationships from the HapMap CEU population release 22. Individuals for association testing in the current analysis were selected from this subset of WGHS participants on the basis of phenotype availability. Association testing for reproductive aging phenotypes was performed with Mach2Qtl v. 1.0.4

b) Replication Studies

The Avon Longitudinal Study of Parents and Children (ALSPAC) mothers
The Avon Longitudinal Study of Parents and Children (ALSPAC) is a prospective study that has been described in detail elsewhere[60] (http://www.alspac.bris.ac.uk). Briefly, 14,541 pregnant women living in one of three Bristol-based health districts in the former County of Avon with an expected delivery date between April 1991 and December 1992 were enrolled in the study. This represented 80–90% of the eligible population. Individuals of known non-white ethnic origin were excluded from all analyses. DNA was collected from mothers as described previously [61]. Genotypes for SNPs and age at menarche data were available in up to 6118 mothers. Mother's recalled age at menarche in completed whole years and adult height were obtained by questionnaires. Ethical approval for the study was obtained from the ALSPAC Law and Ethics Committee and Local Research Ethics Committees.

The Bogalusa Heart Study (BHS):
Between 1973 and 2008, 9 cross-sectional surveys of children aged 4-17 years and 10 cross-sectional surveys of adults aged 18-48 years, who had been previously examined as children, were conducted for CVD risk factor

examinations in Bogalusa, Louisiana. Collection of age at menarche information in the BHS has been previously described[62]. Briefly, girls in the 3rd grade and up were interviewed individually about menstrual history by a registered nurse during the collection of anthropometric measures, health habits, and cardiovascular risk factors. In the ongoing Longitudinal Aging Study funded by NIH and NIA since 2000, there are 1,202 subjects who have been examined 4-14 times from childhood to adulthood and have DNA available for GWA genotyping. Based on the analysis of identity-by-state (IBS) sharing from whole genome genotyping data, we focus on a subset of 343 genotyped women who are of European ancestry, unrelated, and have information about the onset of menarche.

We genotyped 1,202 BHS samples using the Illumina Human610 Genotyping BeadChip[63], and HumanCVD BeadChip[64]. Genotypes were called using a clustering algorithm in Illumina's BeadStudio software. Three samples on the 610 array gave poor results (call rates <99%) and were discarded from the study. In addition, 3 samples had a different estimated gender from genotype data versus gender provided with the phenotype data and were also discarded. SNPs with call rates <90% were discarded, and SNPs with call rates between 90-95% or cluster separation score < 0.3 were manually inspected and cluster positions were edited if needed. We removed approximately 30,000 SNP loci (4.9%) due to poor performance. The final average sample call rate was 99.95% for the 610 BeadChip, and 99.32% for the CVD BeadChip. We assessed reproducibility by genotyping 29 samples in duplicate (18 known replicates, 11 blind replicates), and observed >99.99% identical genotype calls on both BeadChips. Finally we observed 99.98% genotype concordance in 12,581 overlapping SNPs between the 610 and CVD BeadChips. Genotypes were imputed to HapMap release 22 CEU haplotypes using MACH v.1.0.16 (http://www.sph.umich.edu/csg/abecasis/MACH/index.html).

EGCUT (see Stage 1 GWAS study information)

INGI- Carlantino & Friuli Venezia Giulia: Carlantino is a small village in the Province of Foggia in southern Italy that was settled in by founders at the end of sixteenth century. The census in 1,595 counted 10 households in the village. Carlantino has a present-day population of 1,519 inhabitants, and three different surnames account for the majority of living individuals. The endogamy rate, calculated during past century, was 99.5%. Friuli Venezia Giulia: we analysed 4 isolated villages in Northern Italy in the region of Friuli Venezia Giulia. Genotyping was performed using the Illumina INFINIUM 370k CNV array, and imputation was performed using MACH. SNPs were analysed for association with age at menarche using ProbABEL on 322 (Carlantino) and 338 (FVG) women with data on age at menarche.

INGI – Val Borbera: The INGI-Val Borbera population is a collection of 1664 genotyped samples collected in the Val Borbera Valley, a geographically isolated valley located within the Appennine Mountains in NorthWest

Italy[65]. The valley is inhabited by about 3000 descendants from the original inhabitants, living in 7 villages along the valley and in the mountains. The valley was inhabited by about 10,000 people in the 19th century when endogamy was >80% . Around 1930, the population started to decrease due to emigration to South America. Participants were healthy people between 18 and 102 years of age that had at least one grandfather living in the valley. Information on participants was collected during an interview using a standardized medical questionnaire. A total of 910 women declared to have undergone menarche between the age of 9 and 17 (none declared an age at menarche outside this range). Genotyping was performed on an Illumina array 370k Quad v3 and missing data was imputed using MACH. Association testing was conducted using ProbABEL, whilst the variance explained by genetic variants was determined using the GenABEL package within R.

KORA F3 and S4: The KORA study is a series of independent population-based epidemiological surveys of participants living in the region of Augsburg, Southern Germany[66]. All survey participants are residents of German nationality identified through the registration office and were examined in 1994/95 (KORA S3) and 1999/2001 (KORA S4). In the KORA S3 study 4,856 subjects (response 75%), and in KORA S4 in total 4,261 subjects have been examined (response 67%). 3,006 subjects participated in a 10-year follow-up examination of S3 in 2004/05 (KORA F3). For KORA F3 we selected 1,644 subjects of these participants while for KORA S4 we randomly selected 1,814 subjects. Informed consent has been given by all participants. The study has been approved by the local ethics committee.

Genotyping for KORA F3 was performed using Affymetrix 500K Array Set consisting of two chips (Sty I and Nsp I). The KORA S4 samples were genotyped with the Affymetrix Human SNP Array 6.0. Hybridisation of genomic DNA was done in accordance with the manufacturer's standard recommendations. Genotypes were determined using BRLMM clustering algorithm (Affymetrix 500K Array Set) and Birdseed2 clustering algorithm (Affymetrix Array 6.0). For quality control purposes, we applied a positive control and a negative control DNA every 48 (KORA F3) samples or 96 samples (KORA S4). On chip level only subjects with overall genotyping efficiencies of at least 93% were included. In addition the called gender had to agree with the gender in the KORA study database. Imputation of genotypes was performed with the software MACH v1.0.9 (KORA F3) and MACH v1.0.15 (KORA S4) based on HapMap II and analyses were performed in R version 2.8.0.

Orcades: The Orkney Complex Disease Study (ORCADES) is an ongoing family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with the high levels of endogamy historically.
Data for participants from a subgroup of ten islands were used for this analysis. Fasting blood samples were collected and over 200 health-related phenotypes and environmental exposures were measured in each

individual. All participants gave informed consent and the study was approved by Research Ethics Committees in Orkney and Aberdeen.

We genotyped 318,237 SNPs for each individual using the Illumina HumanHap300 beadchip. Alleles were called in BeadStudio using Illumina cluster files. Subjects were excluded if they fulfilled any of the following criteria: genotypic call rate <97%, mismatch between reported and genotypic sex, unexpectedly low genomic sharing with first degree relatives, excess autosomal heterozygosity, or outliers identified by IBS clustering analysis. We excluded SNPs on the basis of minor allele frequency (<0.01), HWE (P<10^-5), call rate (<97%). Pregnant women were excluded from the study. MACH v1.0.15 was used to impute over 2 million SNPs from HapMap build 36. Analyses were implemented using the GenABELand ProbABEL R libraries.

Raine: Recruitment of the Western Australian Pregnancy (Raine) cohort has previously been described in detail[67]. In brief, between 1989 and 1991 2,900 pregnant women were recruited prior to 18-weeks gestation into a randomised controlled trial to evaluate the effects of repeated ultrasound in pregnancy. Recruitment predominantly took place at King Edward Memorial Hospital (Perth, Western Australia). Ninety percent of eligible women agreed to participate in the study. Their 2,868 babies have been followed from recruitment at the average ages of one, two, three, five, eight, ten and 14. Most of the children are of Caucasian ethnicity (82% have two Caucasian parents). Month and year of first menstrual period for girls was recorded. Genotyping was performed using the Illumina 660w quad array and imputation was performed using MACH. Association testing was performed using R (version 2.6.2).

SASBAC: The Swedish and Singapore Breast Association Consortium is a case-control study of breast cancer that has been described previously[68,69]. After exclusions of related individuals and duplicates, ethnic outliers, women without age at menarche and age at menarche outside the range of 9-17 years, data on 723 cases and 685 controls with genome-wide data (call rate>96%) were available. Analyses were performed in SNPtest and STATA 11.0.

SEARCH: The SEARCH ovarian cancer study is an ongoing, population-based ovarian cancer case–control study covering the regions served by the East Anglia and West Midlands cancer registries in the UK and has been described previously[70]. 1126 cases were included in this analysis with genotype data available and age at menarche between 9 and 17 years. Analyses were performed in STATA 10.0.

STR_MZtwins (TWINGENE): Between the years 2004 and 2008 population wide collection of blood on 12,600 twins born 1958 or earlier has been undertaken in a project called TwinGene[71]. The aim of the TwinGene project has been to systematically transform the oldest cohorts of the Swedish Twin Registry (STR) into a

molecular-genetic resource. Beginning in 2004 about 200 twins were contacted each month until the data collection was completed in 2008. When the signed consent forms where returned, the subjects were sent blood sampling equipment and asked to contact a local health facility for blood sampling. Subjects living in vicinity to the cities of Stockholm, Gothenburg, Malmö or Västerås were given the option of visiting hospital blood sampling facilities, in which case the health checkup were omitted. The study population was recruited among twins participating in the Screening Across the Lifespan Twin Study (SALT) which was a telephone interview study conducted in 1998-2002. Other inclusion criteria were that both twins in the pair had to be alive and living in Sweden. Subjects were excluded from the study if they preciously declined participation in future studies or if they had been enrolled in other STR DNA sampling projects. Menopausal information was collected from the SALT interview. In 302 MZ pairs genome-wide array (Illumina 317K) genotyping has been performed. Among these, information on age at menarche was available for at least one of the twins in the pair for 302 pairs. For pairs in which information about menopause were available for both the average within-pair value was used.

VIS, KORCULA and SPLIT: The VIS study, Croatia, is a family-based, cross-sectional study in the isolated island of Vis that included 1,056 examinees aged 18-93. The KORCULA study, Croatia, is a family-based, cross-sectional study in the isolated island of Korcula that included about 965 examinees aged 18-95. The SPLIT study, Croatia, is an ongoing population-based, cross-sectional study in the Dalmatian City of Split that included about 535 examinees aged 18-95. Studies were genotyped using Illumina HAP300v1 (VIS) or Ilumina HAP370CNV (Korcula and SPLIT) and imputation for all studies was performed using MACHv1.16. Analyses were performed using R, GenABEL and ProbABEL.

References for Study Information

1. Agarwala, R., Biesecker, L.G., Hopkins, K.A., Francomano, C.A. & Schaffer, A.A. Software for constructing and verifying pedigrees within large genealogies and an application to the Old Order Amish of Lancaster County. *Genome Res* **8**, 211-21 (1998).
2. Beiler, K. *Church Directory of the Lancaster County Amish*, (Peqaea Publishers, 1996).
3. Rampersaud, E. et al. The association of coronary artery calcification and carotid artery intima-media thickness with distinct, traditional coronary artery disease risk factors in asymptomatic adults. *Am J Epidemiol* **168**, 1016-23 (2008).
4. Sorkin, J. et al. Exploring the genetics of longevity in the Old Order Amish. *Mech Ageing Dev* **126**, 347-50 (2005).
5. Mitchell, B.D. et al. The genetic response to short-term interventions affecting cardiovascular function: rationale and design of the Heredity and Phenotype Intervention (HAPI) Heart Study. *Am Heart J* **155**, 823-8 (2008).

6. Shuldiner, A.R. et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *JAMA* **302**, 849-57 (2009).
7. Streeten, E.A. et al. Reduced incidence of hip fracture in the Old Order Amish. *J Bone Miner Res* **19**, 308-13 (2004).
8. The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* **129**, 687-702 (1989).
9. Stark, O., Peckham, C.S. & Moynihan, C. Weight and age at menarche. *Arch Dis Child* **64**, 383-7 (1989).
10. Barrett, J.C. et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* **41**, 703-707 (2009).
11. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).
12. Power, C. & Elliott, J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol* **35**, 34-41 (2006).
13. Firmann, M. et al. The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovasc Disord* **8**, 6 (2008).
14. Olsen, J. et al. The Danish National Birth Cohort--its background, structure and aim. *Scand J Public Health* **29**, 300-7 (2001).
15. Metspalu, A. The Estonian Genome Project. *Drug Dev Res* **62**, 97-101 (2004).
16. Nelis, M. et al. Genetic structure of Europeans: a view from the North-East. *PLoS One* **4**, e5472 (2009).
17. Day, N. et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* **80 Suppl 1**, 95-103 (1999).
18. Aulchenko, Y.S. et al. Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* **12**, 527-34 (2004).
19. Dawber, T.R., Meadors, G.F. & Moore, F.E., Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health* **41**, 279-81 (1951).
20. Dawber, T.R., Kannel, W.B. & Lyell, L.P. An approach to longitudinal studies in a community: the Framingham Study. *Ann N Y Acad Sci* **107**, 539-56 (1963).
21. Feinleib, M., Kannel, W.B., Garrison, R.J., McNamara, P.M. & Castelli, W.P. The Framingham Offspring Study. Design and preliminary data. *Prev Med* **4**, 518-25 (1975).
22. Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J. & Castelli, W.P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* **110**, 281-90 (1979).

23. Splansky, G.L. et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* **165**, 1328-35 (2007).
24. Yliharsila, H. et al. Body mass index during childhood and adult body composition in men and women aged 56-70 y. *Am J Clin Nutr* **87**, 1769-75 (2008).
25. Heistaro, S. Methodology report. Health 2000 survey. Publications of National Public Health Institute. in *Series B 26* (2008).
26. Ferrucci, L. et al. Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *J Am Geriatr Soc* **48**, 1618-25 (2000).
27. Epstein, M.P., Duren, W.L. & Boehnke, M. Improved inference of relationship for pairs of individuals. *Am J Hum Genet* **67**, 1219-31 (2000).
28. Koller, D.L. et al. Genome-wide association study of bone mineral density in premenopausal European-American women and replication in African-American women. *J Clin Endocrinol Metab* **95**, 1802-9 (2010).
29. Colditz, G.A. & Hankinson, S.E. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* **5**, 388-96 (2005).
30. Hunter, D.J. et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* **39**, 870-4 (2007).
31. Cornelis, M.C. et al. Joint effects of common genetic variants on the risk for type 2 diabetes in U.S. men and women of European ancestry. *Ann Intern Med* **150**, 541-50 (2009).
32. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. *Diabetes* **28**, 1039-57 (1979).
33. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* **23 Suppl 1**, S4-19 (2000).
34. Hu, F.B. et al. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med* **161**, 1542-8 (2001).
35. Korn, J.M. et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet* **40**, 1253-60 (2008).
36. Taponen, S. et al. Metabolic cardiovascular disease risk factors in women with self-reported symptoms of oligomenorrhea and/or hirsutism: Northern Finland Birth Cohort 1966 Study. *J Clin Endocrinol Metab* **89**, 2114-8 (2004).
37. Laitinen, J., Power, C. & Jarvelin, M.R. Family social class, maternal body mass index, childhood body mass index, and age at menarche as predictors of adult obesity. *Am J Clin Nutr* **74**, 287-94 (2001).
38. Sabatti, C. et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* **41**, 35-46 (2009).

39. Boomsma, D.I. et al. Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* **9**, 849-57 (2006).
40. Boomsma, D.I. et al. Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. *Eur J Hum Genet* **16**, 335-42 (2008).
41. Zhu, G. et al. A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs. *Am J Hum Genet* **65**, 483-92 (1999).
42. Wright, M.J. & Martin, N.G. Brisbane Adolescent Twin Study: outline of study methods and research projects. *Aust J Psychol* **56**, 65-78 (2004).
43. Anderson, C.A. et al. A genome-wide linkage scan for age at menarche in three populations of European descent. *J Clin Endocrinol Metab* **93**, 3965-70 (2008).
44. Medland, S.E. et al. Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet* **85**, 750-5 (2009).
45. Hofman, A. et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol* **22**, 819-29 (2007).
46. Estrada, K. et al. GRIMP: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics* **25**, 2750-2 (2009).
47. Bucholz, K.K. et al. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *J Stud Alcohol* **55**, 149-58 (1994).
48. Gunderson, K.L. et al. Whole-genome genotyping. *Methods Enzymol* **410**, 359-76 (2006).
49. Pilia, G. et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* **2**, e132 (2006).
50. Willer, C.J. et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* **40**, 161-9 (2008).
51. Chen, W.M. & Abecasis, G.R. Family-based association tests for genomewide association scans. *Am J Hum Genet* **81**, 913-26 (2007).
52. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. *Annu Rev Genomics Hum Genet* **10**, 387-406 (2009).
53. Burdick, J.T., Chen, W.M., Abecasis, G.R. & Cheung, V.G. In silico method for inferring genotypes in pedigrees. *Nat Genet* **38**, 1002-4 (2006).
54. Andrew, T. et al. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Res* **4**, 464-77 (2001).
55. Spector, T.D. & Williams, F.M. The UK Adult Twin Registry (TwinsUK). *Twin Res Hum Genet* **9**, 899-906 (2006).

56. Spector, T.D. & MacGregor, A.J. The St. Thomas' UK Adult Twin Registry. *Twin Res* **5**, 440-3 (2002).
57. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).
58. Ridker, P.M. et al. Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin Chem* **54**, 249-55 (2008).
59. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
60. Golding, J., Pembrey, M. & Jones, R. ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol* **15**, 74-87 (2001).
61. Jones, R., W. et al. A DNA bank established as part of the Avon longitudinal study of pregnancy and childhood (ALSPAC). *European Journal of Human Genetics* **8**, 653-660 (2000).
62. Wattigney, W.A., Srinivasan, S.R., Chen, W., Greenlund, K.J. & Berenson, G.S. Secular trend of earlier onset of menarche with increasing obesity in black and white girls: the Bogalusa Heart Study. *Ethn Dis* **9**, 181-9 (1999).
63. Eberle, M.A. et al. Power to detect risk alleles using genome-wide tag SNP panels. *PLoS Genet* **3**, 1827-37 (2007).
64. Keating, B.J. et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS One* **3**, e3583 (2008).
65. Traglia, M. et al. Heritability and demographic analyses in the large isolated population of Val Borbera suggest advantages in mapping complex traits genes. *PLoS One* **4**, e7554 (2009).
66. Wichmann, H.E., Gieger, C. & Illig, T. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1**, S26-30 (2005).
67. Newnham, J.P., Evans, S.F., Michael, C.A., Stanley, F.J. & Landau, L.I. Effects of frequent ultrasound during pregnancy: a randomised controlled trial. *Lancet* **342**, 887-91 (1993).
68. Magnusson, C. et al. Breast-cancer risk following long-term oestrogen- and oestrogen-progestin-replacement therapy. *Int J Cancer* **81**, 339-44 (1999).
69. Einarsdottir, K. et al. Comprehensive analysis of the ATM, CHEK2 and ERBB2 genes in relation to breast tumour characteristics and survival: a population-based case-control and follow-up study. *Breast Cancer Res* **8**, R67 (2006).
70. Song, H. et al. Common variants in mismatch repair genes and risk of invasive ovarian cancer. *Carcinogenesis* **27**, 2235-42 (2006).
71. Rahman, I. et al. Genetic dominance influences blood biomarker levels in a sample of 12,000 Swedish elderly twins. *Twin Res Hum Genet* **12**, 286-94 (2009).

EXTENDED ACKNOWLEDGEMENTS

AGES: The Age, Gene/Environment Susceptibility-Reykjavik Study is funded by NIH contract N01-AG-12100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). Genotyping was conducted at the NIA IRP Laboratory of Neurogenetics.

Amish: The Amish study was supported by NIH research grants U01 HL72515, U01 HL84756 and R01 088119, the University of Maryland General Clinical Research Center, grant M01 RR 16500, the Clinical Nutrition Research Unit of Maryland (P30 DK072488), and the Baltimore Veterans Administration Medical Center Geriatrics Research

ARIC: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367, R01HL086694 and RC2 HL102419; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

B58C: We acknowledge use of genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. B58C-T1DGC (British 1958 birth cohort Type 1 Diabetes Genetic Consortium controls) analyses utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. T1DGC GWAS data were deposited by the Diabetes & Inflammation Laboratory, Cambridge Institute for Medical Research, University of Cambridge (John Todd, Helen Stevens and Neil Walker) which is funded by the Juvenile Diabetes Research Foundation International, The National Institute for Health Research Cambridge Biomedical Research Centre and the Wellcome Trust (Strategic Award 079895).

CoLaus: The CoLaus study was supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, Switzerland, and the Swiss National Science Foundation (Ref: 33CSCO-122661). The

collaboration between CoLaus and the Medical Research Council was supported by a specific project grant from the Medical Research Council (Ref: G0701863). SB is supported by the Giorgi-Cavaglieri Foundation, the Swiss National Science Foundation (grant 3100AO-116323/1), the Leenaards Foundation, the European Union HEALTH-F4-2007-201550 HYPERGENES grant and the Swiss Institute of Bioinformatics.

DNBC: The Danish National Research Foundation established the Danish Epidemiology Science Centre, which initiated and created the Danish National Birth Cohort. The cohort was established with the support of a major grant from this foundation. Additional support for the Danish National Birth Cohort has been obtained from the Danish Pharmacists' Fund, the Egmont Foundation, the March of Dimes Birth Defects Foundation, the Augustinus Foundation and the Health Fund of the Danish Health Insurance Societies. The generation of GWAS genotype data for the DNBC samples was carried out within the GENEVA consortium with funding provided through the NIH Genes, Environment and Health Initiative (GEI) (U01HG004423). Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01HG004446). Genotyping was performed at Johns Hopkins University Center for Inherited Disease Research, with support from the NIH GEI (U01HG004438).

EGCUT: EGCUT authors received support from FP7 grants (201413 ENGAGE, 212111 BBMRI, 205419 ECOGENE, 245536 OPENGENE). EGCUT authors also received targeted financing from Estonian Government SF0180142s08 and by EU via the European Regional Development Fund, in the frame of Centre of Excellence in Genomics. The genotyping of the Estonian Genome Project samples were performed in Estonian Biocentre Genotyping Core Facility, EGCUT authors want to acknowledge Viljo Soo for his contribution in that. EGC data analyzes were carried out in part in the High Performance Computing Center of University of Tartu.

EPIC-Norfolk: The EPIC Norfolk Study is funded by Cancer Research United Kingdom and the Medical Research Council. I.B. acknowledges support from the Wellcome Trust (Ref: 077016/Z/05/Z) and EU FP6 funding (contract no LSHM-CT-2003-503041). S.L. is supported by a grant from Unilever Corporate Research.

Erasmus Rucphen Family (ERF): The study was supported by grants from The Netherlands Organisation for Scientific Research (NWO), Erasmus MC and the Centre for Medical Systems Biology (CMSB). We are grateful to all general practitioners for their contributions, to Petra Veraart for her help in genealogy, Jeannette Vergeer for the supervision of the laboratory work and Peter Snijders for his help in data collection.

The **Framingham Heart Study** phenotype-genotype analyses and the meta-analysis for this work were supported by the National Institute of Aging (Genetics of Reproductive Life Period and Health Outcomes,

R21AG032598; JMM, KL, DEK, DPK) and the National Institute of Arthritis Musculoskeletal and Skin Diseases and National Institute on Aging R01 AR/AG 41398 (DPK). The Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study Contract No. N01-HC-25195 and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). Analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. A portion of this research was conducted using the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.

HBCS, Health 2000, NFBC: The three Finnish studies were funded/partly funded through the European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project, grant agreement HEALTH-F4-2007- 201413, by the Wellcome Trust (grant n.o. 89061/Z/09/Z), by the European Commission under the programme 'Quality of Life and Management of the Living Resources' of 5th Framework Programme (no. QLG2-CT-2002-01254), by the Academy of Finland (Finnish Centre of Excellence in Complex Disease Genetics - grant n.o. 129680, grant n.o. 120315 and 129287 to EW, and grant no. 129494 to VS), by the Medical Research Council (studentship grant G0500539 to US) and by the National Heart, Lung, and Blood Institute grant HL087679 through the STAMPEED program.

InCHIANTI: The InCHIANTI study baseline (1998-2000) was supported as a “targeted project” (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336); the InCHIANTI Follow-up 1 (2001-2003) was funded by the U.S. National Institute on Aging (Contracts: N.1-AG-1-1 and N.1-AG-1-2111); the InCHIANTI Follow-ups 2 and 3 studies (2004-2010) were financed by the U.S. National Institute on Aging (Contract: N01-AG-5-0002);supported in part by the Intramural research program of the National Institute on Aging, National Institutes of Health, Baltimore, Maryland.

Indiana: We thank the individuals who participated in this study, as well as the study coordinators, without whom this work would not have been possible. This work was supported by National Institutes of Health grants P01 AG-18397 and M01 RR-00750. Genotyping services were provided by CIDR. CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Library of Medicine.

NHS: The NHS GWA scan was performed as part of the Cancer Genetic Markers of Susceptibility initiative of the NCI. We particularly acknowledge the contributions of R. Hoover, A. Hutchinson, K. Jacobs and G. Thomas. The current research is supported by CA 40356 and U01-CA98233 from the National Cancer Institute.

The NHS/HPFS type 2 diabetes GWAS (U01HG004399) is a component of a collaborative project that includes 13 other GWAS funded as part of the Gene Environment-Association Studies (GENEVA) under the NIH Genes, Environment and Health Initiative (GEI) (U01HG004738, U01HG004422, U01HG004402, U01HG004729, U01HG004726, U01HG004735, U01HG004415, U01HG004436, U01HG004423, U01HG004728, RFAHG006033) with additional support from individual NIH Institutes (NIDCR: U01DE018993, U01DE018903; NIAAA: U10AA008401, NIDA: P01CA089392, R01DA013423; NCI: CA63464, CA54281, CA136792, Z01CP010200). Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Genotyping was performed at the Broad Institute of MIT and Harvard, with funding support from the NIH GEI (U01HG04424), and Johns Hopkins University Center for Inherited Disease Research, with support from the NIH GEI (U01HG004438) and the NIH contract "High throughput genotyping for studying the genetic contributions to human disease"(HHSN268200782096C). Additional funding for the current research was provided by the National Cancer Institute (NCI, P01CA087969, P01CA055075), and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK, R01DK058845).

We acknowledge the study participants in the NHS for their contribution in making this study possible.

NTR: Financial support was obtained from the Netherlands Organization for Scientific Research (NWO): Database Twin Register (575-25-006), Twin-family database for behavior genetics and genomic studies (480-04-004); Resolving cause and effect in the association between exercise and well-being (904-61-193); Center for Medical Systems Biology (NWO Genomics); Spinozapremie (56-464-14192); Centre for Neurogenomics and Cognitive Research (CNCR-VU); Genome wide analyses of European twin and population cohorts (EU/QLRT-2001-01254); Genetics of Mental Illness (ERC-230374). Genotyping was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health.

QIMR (Queensland Institute of Medical Research)**:** We thank the twins and their families for their participation. We also thank Dixie Statham, Ann Eldridge, Marlene Grace, Kerrie McAloney (sample collection); Lisa Bowdler, Steven Crooks (DNA processing); Sarah Medland, Dale Nyholt and Scott Gordon (Imputation and genotyping QC). Funding was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498), the

Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921), the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254), and the U.S. National Institutes of Health (NIH grants AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, MH66206). A portion of the genotyping on which this study was based (Illumina 370K scans on 4300 individuals) was carried out at the Center for Inherited Disease Research, Baltimore (CIDR), through an access award to our late colleague Dr. Richard Todd (Psychiatry, Washington University School of Medicine, St Louis). Statistical analyses were carried out on the Genetic Cluster Computer, which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003). E.M.B is supported by University of Queensland IPRS and UQILAS awards. G.W.M. is supported by the National Health and Medical Research Council (NHMRC) Fellowship Scheme.

Rotterdam Studies: The generation and management of GWAS genotype data for the **Rotterdam Study** is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI) - Netherlands Consortium of Healthy Aging (NCHA) project nr. 050-060-810, and funding from the European Commision (HEALTH-F2-2008-201865, GEFOS; HEALTH-F2-2008-35627, TREAT-OA). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. We thank Pascal Arp, Mila Jhamai, Dr Michael Moorhouse, Marijn Verkerk, and Sander Bervoets for their help in creating the GWAS database. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practioners and pharmacists. We would like to thank Dr. Tobias A. Knoch, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technology under grants #01 AK 803 A-H and # 01 IG 07015 G for access to their grid resources.

SAGE: Funding support for the Study of Addiction: Genetics and Environment
(SAGE) was provided through the NIH Genes, Environment and Health Initiative [GEI] (U01 HG004422). SAGE is one of the genome-wide association studies funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Support for collection of datasets and samples was provided by the Collaborative Study on the Genetics of Alcoholism (COGA; U10 AA008401), the Collaborative Genetic Study of Nicotine Dependence (COGEND; P01 CA089392), and the Family Study of

Cocaine Dependence (FSCD; R01 DA013423). Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by the NIH GEI (U01HG004438), the National Institute on Alcohol Abuse and Alcoholism, the National Institute on Drug Abuse, and the NIH contract "High throughput genotyping for studying the genetic contributions to human disease" (HHSN268200782096C).

The Collaborative Study on the Genetics of Alcoholism (COGA), Principal Investigators B. Porjesz, V. Hesselbrock, H. Edenberg, L. Bierut, includes ten different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, J. Nurnberger Jr., T. Foroud); University of Iowa (S. Kuperman, J. Kramer); SUNY Downstate (B. Porjesz); Washington University in St. Louis (L. Bierut, A. Goate, J. Rice, K. Bucholz); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield); Southwest Foundation (L. Almasy), Howard University (R. Taylor) and Virginia Commonwealth University (D. Dick). Other COGA collaborators include: L. Bauer University of Connecticut); D. Koller, S. O'Connor, L. Wetherill, X. Xuei (Indiana University); Grace Chan (University of Iowa); N. Manz, M. Rangaswamy (SUNY Downstate); A. Hinrichs, J. Rohrbaugh, J-C Wang (Washington University in St. Louis); A. Brooks (Rutgers University); and F. Aliev (Virginia Commonwealth University). A. Parsian and M. Reilly are the NIAAA Staff Collaborators. We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding PI and Co-PI of COGA, and also owe a debt of gratitude to other past organizers of COGA, including Ting-Kai Li, currently a consultant with COGA, P. Michael Conneally, Raymond Crowe, and Wendy Reich, for their critical contributions. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA).

The Collaborative Genetic Study of Nicotine Dependence (COGEND) project is a collaborative research group and part of the NIDA Genetics Consortium. Subject collection was supported by NIH grant P01 CA89392 (PI - L Bierut) from the National Cancer Institute. Lead investigators directing data collection are Laura Bierut, Naomi Breslau, Dorothy Hatsukami, and Eric Johnson. In memory of Theodore Reich, founding Principal Investigator of COGEND, we are indebted to his leadership in the establishment and nurturing of COGEND and acknowledge with great admiration his seminal scientific contributions to the field. The authors thank Heidi Kromrei and Tracey Richmond for their assistance in data collection.

The **SardiNIA** study is supported by the National Institute on Aging, NIH, contract NO1-AG-1-2109 to the SardiNIA ("ProgeNIA") team. The authors are grateful to all of the volunteers who participated in this study, Monsignore Piseddu, Bishop of Ogliastra, the mayors and citizens of the Sardinian towns (Lanusei, Ilbono, Arzana, and Elini), the head of the Public Health Unit ASL4 for their volunteerism and cooperation, and team of physicians, nurses, biologists and the recruitment personnel. This research was supported in part by the

Intramural Research Program of the NIH, National Institute on Aging. We thank Osorio Meirelles for careful reading and comments on the manuscript.

Twins UK (TUK): The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F2-2008-201865-GEFOS and (FP7/2007-2013), ENGAGE project grant agreement HEALTH-F4-2007-201413 and the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from the Dept of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. TDS is an NIHR senior Investigator. The project also received support from a *Biotechnology and Biological Sciences Research Council* (BBSRC) project grant. (G20234) .The authors acknowledge the funding and support of the National Eye Institute via an NIH/CIDR genotyping project (PI: Terri Young) ". We thank the staff from the Genotyping Facilities at the Wellcome Trust Sanger Institute for sample preparation, Quality Control and Genotyping led by Leena Peltonen and Panos Deloukas; Le Centre National de Génotypage, France, led by Mark Lathrop, for genotyping; Duke University, North Carolina, USA, led by David Goldstein, for genotyping; and the Finnish Institute of Molecular Medicine, Finnish Genome Center, University of Helsinki, led by Aarno Palotie. Genotyping was also performed by CIDR as part of an NEI/NIH project grant.

WGHS: The WGHS is supported by HL 043851 and HL69757 from the National Heart, Lung, and Blood Institute and CA 047988 from the National Cancer Institute, the Donald W. Reynolds Foundation and the Fondation Leducq, with collaborative scientific support and funding for genotyping provided by Amgen.

Replication Studies:

ALSPAC: We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council, the Wellcome Trust (including grant number 076467/Z/05/Z) and the University of Bristol provide core support for ALSPAC. Repeated questionnaire assessment of pubertal status was funded by the Centers for Disease Control and Prevention, USA.

BHS:(ENS = Erin N. Smith, NJS = Nicholas J. Schork, SSM = Sarah S. Murray, WC = Wei Chen, SRS = Sathanur R. Srinivasan, and GSB = Gerald S. Berenson.)
ENS, NJS, and SSM are supported in part by National Institutes of Health grant NIH 1U54RR025204-01 and UL1 RR025774. WC, SRS and GSB are supported by grants 0855082E from American Heart Association, HD-061437

from the National Institute of Child Health and Human Development, and AG-16592 from the National Institute on Aging.

INGI- Carlantino & Friuli Venezia Giulia
The authors would like to thank the populations of the villages contributing to these analyses as well as grants from the FVG region and Fondo Treiste.

INGI – Val Borbera: We thank the inhabitants of the villages in Val Borbera who have made this study possible. We thank the Val Borbera municipalities, the Tortona and Genova archidiocese and the ASL 22 of Novi Ligure (Alessandria) for support. We are indebted to Mrs Maria Rosa Biglieri for help in the clinical evaluation of the population. We acknowledge the collaboration of the MDs from the School of Internal Medicine and Cardiology of the San Raffaele Hospital and of the local doctors. We especially thank Dr. Diego Sabbi for many suggestions and great support. The research was funded by grants of the Italian Health Ministry (RF-FSR-2007-647201) by Fondazione Compagnia di San Paolo, Fondazione Cassa di Risparmio di Alessandria and Telethon Foundation Onlus Rome.

Health Ministry project RF-FSR-2007-647201, Fondazione Compagnia di San Paolo. Fondazione Cassa di Risparmio di Alessandria and Telethon Foundation Onlus Rome.

KORA F3 and S4: The KORA research platform was initiated and financed by the Helmholtz Center Munich, German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Part of this work was financed by the German National Genome Research Network (NGFN-2 and NGFNPlus: 01GS0823). Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ.

Orcades: ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

Raine: The authors are grateful to the Raine Foundation, to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the NH&MRC for their long term contribution to funding the study over the last 20 years and also the following Institutions for providing funding for Core Management of the Raine Study:

o The University of Western Australia (UWA)
o Raine Medical Research Foundation
o UWA Faculty of Medicine, Dentistry and Health Sciences
o The Telethon Institute for Child Health Research
o Women and Infants Research Foundation
The authors gratefully acknowledge the assistance of the Western Australian Genetic Epidemiology Resource and the Western Australian DNA Bank (both National Health and Medical Research Council of Australia National Enabling Facilities). The authors also acknowledge the support of the National Health and Medical Research Council of Australia (Grant ID 572613 and ID 003209) and the Canadian Institutes of Health Research (Grant ID 166067). We gratefully acknowledge the assistance of the Wind Over Water Foundation, the Telethon Institute for Child Health Research, and the Raine Medical Research Foundation of the University of Western Australia.

SASBAC: The SASBAC study was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the United States National Institute of Health (NIH CA104021) and the Susan G. Komen Breast Cancer Foundation. E.I. was supported by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Society of Medicine, the Swedish Foundation for Strategic Research, and the Royal Swedish Academy of Science while working with this article.

STR_MZtwins (TwinGene)
The TWINGENE study was supported by the Swedish Research Council, GenomEUtwin, The United States National Institute of Health (NIH), Swedish Foundation for Strategic Research (SSF) and the Swedish Ministry of Higher Education and Research.

VIS, KORCULA and SPLIT: The VIS, KORCULA and SPLIT studies were funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health. The SNP genotyping for the VIS cohort was performed in the core genotyping laboratory of the Wellcome Trust Clinical Research Facility at the Western General Hospital, Edinburgh, Scotland. The SNP genotyping for the KORCULA cohort was performed in Helmholtz Zentrum München, Neuherberg, Germany. The SNP genotyping for the SPLIT cohort was performed by AROS Applied Biotechnology, Aarhus, Denmark.

Members of the GIANT (Genetic Investigation of Anthropometric Traits) Consortium are:

1) BMI group

Cristen J. Willer[1], Elizabeth K. Speliotes[2,3], Ruth J.F. Loos[4,5], Shengxu Li[4,5], Cecilia M. Lindgren[6], Iris M. Heid[7], Sonja I. Berndt[8], Amanda L. Elliott[9,10], Anne U. Jackson[1], Claudia Lamina[7], Guillaume Lettre[9,11], Noha Lim[12], Helen N. Lyon[3,11], Steven A. McCarroll[9,10], Konstantinos Papadakis[13], Lu Qi[14,15], Joshua C. Randall[6], Rosa Maria Roccasecca[16], Serena Sanna[17], Paul Scheet[18], Michael N. Weedon[19], Eleanor Wheeler[16], Jing Hua Zhao[4,5], Leonie C. Jacobs[20], Inga Prokopenko[6,21], Nicole Soranzo[16,22], Toshiko Tanaka[23], Nicholas J. Timpson[24], Peter Almgren[25], Amanda Bennett[26], Richard N. Bergman[27], Sheila A. Bingham[28,29], Lori L. Bonnycastle[30], Morris Brown[31], Noël P. Burtt[9], Peter Chines[30], Lachlan Coin[32], Francis S. Collins[30], John M. Connell[33], Cyrus Cooper[34], George Davey Smith[24], Elaine M. Dennison[34], Parimal Deodhar[30], Paul Elliott[32], Michael R. Erdos[30], Karol Estrada[20], David M. Evans[24], Lauren Gianniny[9], Christian Gieger[7], Christopher J. Gillson[4,5], Candace Guiducci[9], Rachel Hackett[9], David Hadley[13], Alistair S. Hall[35], Aki S. Havulinna[36], Johannes Hebebrand[37], Albert Hofman[38], Bo Isomaa[39], Kevin B. Jacobs[40], Toby Johnson[41,42,43], Pekka Jousilahti[36], Zorica Jovanovic[5,44], Kay-Tee Khaw[45], Peter Kraft[46], Mikko Kuokkanen[9,47], Johanna Kuusisto[48], Jaana Laitinen[49], Edward G. Lakatta[50], Jian'an Luan[4,5], Robert N. Luben[45], Massimo Mangino[51], Wendy L. McArdle[52], Thomas Meitinger[53,54], Antonella Mulas[17], Patricia B. Munroe[55], Narisu Narisu[30], Andrew R. Ness[56], Kate Northstone[52], Stephen O'Rahilly[5,44], Carolin Purmann[5,44], Matthew G. Rees[30], Martin Ridderstråle[57], Susan M. Ring[52], Fernando Rivadeneira[20,38], Aimo Ruokonen[58], Manjinder S. Sandhu[4,45], Jouko Saramies[59], Laura J. Scott[1], Angelo Scuteri[60], Kaisa Silander[47], Matthew A. Sims[4,5], Kijoung Song[12], Jonathan Stephens[61], Suzanne Stevens[51], Heather M. Stringham[1], Y.C. Loraine Tung[5,44], Timo T. Valle[62], Cornelia M. Van Duijn[38], Karani S. Vimaleswaran[4,5], Peter Vollenweider[63], Gerard Waeber[63], Chris Wallace[55], Richard M. Watanabe[64], Dawn M. Waterworth[12], Nicholas Watkins[61], The Wellcome Trust Case Control Consortium[65], Jacqueline C.M. Witteman[38], Eleftheria Zeggini[6], Guangju Zhai[22], M. Carola Zillikens[20], David Altshuler[9,10], Mark J. Caulfield[55], Stephen J. Chanock[8], I. Sadaf Farooqi[5,44], Luigi Ferrucci[23], Jack M. Guralnik[66], Andrew T. Hattersley[67], Frank B. Hu[14,15], Marjo-Riitta Jarvelin[32], Markku Laakso[48], Vincent Mooser[12], Ken K. Ong[4,5], Willem H. Ouwehand[16,61], Veikko Salomaa[36], Nilesh J. Samani[51], Timothy D. Spector[22], Tiinamaija Tuomi[68,69], Jaakko Tuomilehto[62], Manuela Uda[17], André G. Uitterlinden[20,38], Nicholas J. Wareham[4,5], Panagiotis Deloukas[16], Timothy M. Frayling[19], Leif C. Groop[25,70], Richard B. Hayes[8], David J. Hunter[9,14,15,46], Karen L. Mohlke[71], Leena Peltonen[9,16,72], David Schlessinger[73], David P. Strachan[13], H-Erich Wichmann[7,74], Mark I. McCarthy[6,21,75], Michael Boehnke[1], Inês Barroso[16], Gonçalo R. Abecasis[18], Joel N. Hirschhorn[3,11,76]

1. Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA
2. Division of Gastroenterology, Massachusetts General Hospital, Boston, MA 02114, USA
3. Metabolism Initiative and Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Boston, MA 02142, USA
4. MRC Epidemiology Unit, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK
5. Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK
6. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK
7. Institute of Epidemiology, Helmholtz Zentrum München, Ingolstaedter Landstr. 1, 85764 Neuherberg,

Germany
8. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD 20892, USA
9. Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
10. Department of Molecular Biology, Massachusetts General Hospital, Cambridge, MA 02144, USA
11. Program in Genomics and Divisions of Endocrinology and Genetics, Children's Hospital, Boston, MA 02115, USA
12. Medical Genetics/Clinical Pharmacology and Discovery Medicine, PA 19406, USA
13. Division of Community Health Sciences, St George's, University of London, London SW17 0RE, UK

14. Department of Nutrition, Harvard School of Public Health, Boston, MA 02115, USA
15. Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA
16. Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK
17. Istituto di Neurogenetica e Neurofarmacologia, Consiglio Nazionale delle Ricerche, Cagliari, Italy
18. Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA
19. Genetics of Complex Traits, Peninsula Medical School, Exeter EX1 2LU, UK
20. Department of Internal Medicine, Erasmus MC, PO Box 2400, NL-3000-CA Rotterdam, The Netherlands
21. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford OX3 7LJ, UK
22. Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK
23. National institute of Aging, Clinical research branch - Longitudinal Studies Section, Baltimore, MD 21225, USA
24. MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, Bristol BS8 2PR, UK
25. Lund University Diabetes Centre, Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden
26. DRL, OCDEM, Churchill Hospital, Headington, Oxford OX3 7LJ, UK
27. Physiology and Biophysics, University of Southern California School of Medicine, Los Angeles, CA 90033, USA
28. MRC Dunn Human Nutrition Unit, Wellcome Trust/MRC Building, Cambridge CB2 0XY, UK
29. MRC Centre for Nutritional Epidemiology in Cancer Prevention and Survival, Cambridge CB1 8RN, UK
30. National Human Genome Research Institute, Bethesda, MD 20892, USA
31. Clinical Pharmacology Unit, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK
32. Department of Epidemiology and Public Health, Imperial College London, St Mary's Campus, Norfolk Place, W2 1PG London, UK
33. BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, University of Glasgow, Glasgow G12 8TA, UK
34. MRC Epidemiology Resource Centre, (University of Southampton), Southampton General Hospital, Southampton SO16 6YD, UK.
35. Yorkshire Heart Centre, Leeds General Infirmary, Leeds LS1 3EX, UK
36. KTL-National Public Health Institute, FI-00300 Helsinki, Finland
37. Department of Child and Adolescent Psychiatry, University of Duisburg-Essen, Virchowstr. 174, 45147 Essen, Germany
38. Department of Epidemiology, Erasmus MC, PO Box 2400, NL-3000-CA Rotterdam, The Netherlands
39. Folkhalsan Research Center, Malmska Municipal Health Center and Hospital, Jakobstad, Finland
40. Bioinformed Consulting Services, Gaithersburg, MD 20877, USA
41. Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland
42. University Institute for Social and Preventative Medicine, Centre Hospitalier Universitaire Vaudois (CHUV), 1005 Lausanne, Switzerland
43. Swiss Institute of Bioinformatics, Switzerland
44. University of Cambridge Metabolic Research Laboratories, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK
45. Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge CB2 0SR, UK
46. Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
47. Department of Molecular Medicine, National Public Health Institute, FIN-00300 Helsinki, Finland
48. Department of Medicine, University of Kuopio, 70210 Kuopio, Finland
49. Finnish Institute of Occpational Health, Aapistie 1, Fin-90220 Oulu, Finland
50. Laboratory of Cardiovascular Science, Gerontology Research Center, National Institute on Aging,Baltimore, MD 21224, USA
51. Department of Cardiovascular Sciences, University of Leicester, Clinical Sciences, Glenfield General Hospital, Leicester LE3 9QP, UK
52. ALSPAC, Department of Social Medicine, University of Bristol, Bristol BS8 1TQ, UK
53. Institute of Human Genetics, Helmholtz Zentrum München, Ingolstaedter Landstr. 1, 85764 Neuherberg, Germany

54. Institute of Human Genetics, Technical University Munich, D-81765, Munich, Germany
55. Clinical Pharmacology, The William Harvey Research Institute, Bart's and The London, Queen Mary's School of Medicine and Dentistry, Charterhouse Square, London EC1M 6BQ, UK

56. Department of Oral & Dental Science, Bristol BS1 2LY, UK
57. Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden
58. Department of Clinical Chemistry, University of Oulu, Fin-90220, Oulu, Finland
59. Savitaipale Health Center, Savitaipale, Finland
60. Unita' Operativa Geriatria - Istituto Nazionale Ricovero e Cura Anziani - Rome, Italy
61. Department of Haematology, University of Cambridge/NHS Blood & Transplant, Cambridge CB2 2PR, UK
62. National Public Health Institute, Department of Epidemiology and Health Promotion, Mannerheimintie 166, FIN-00300 Helsinki, FINLAND
63. Department of Internal Medicine, BH-10 Centre Hospitalier Universitaire Vaudois (CHUV), 1011 Lausanne, Switzerland
64. Department of Preventive Medicine, Division of Biostatistics, Keck School of Medicine, University of Southern California, CHP-220, Los Angeles, CA 90089, USA
65. The membership of this consortium is listed in the supplementary material.
66. Laboratory of Epidemiology, Demography, and Biometry; Gerontology Research Center, National Institute on Aging, Bethesda, MD 20892, USA
67. Peninsula Medical School, Exeter EX5 2DW, UK
68. 3 Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland
69. Research Program of Molecular Medicine, University of Helsinki, Helsinki, Finland
70. Department of Medicine, Helsinki University, Helsinki, Finland
71. Department of Genetics, University of North Carolina, CB #7264, Chapel Hill, NC 27599, USA
72. Institute of Molecular Medicine, University of Helsinki, Finland
73. Laboratory of Genetics, NIH Biomedical Research Center, National Institute on Aging, Baltimore, MD 21224, USA
74. Institute of Medical Information Processing, Biometry, and Epidemiology, Ludwig-Maximilians-University München, Marchioninistr. 15, 81377 München, Germany
75. NIHR Oxford Biomedical Research Centre, University of Oxford, Old Road, Headington, Oxford OX3 7LJ, UK
76. Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

2) Height group

Lango Allen,Hana[1], Estrada,Karol[2,3,4], Lettre,Guillaume[5,6], Berndt,Sonja I.[7], Weedon,Michael N.[1], Rivadeneira,Fernando[2,3,4], Willer,Cristen J.[8], Jackson,Anne U.[8], Vedantam,Sailaja[9,10], Raychaudhuri,Soumya[11,12], Ferreira,Teresa[13], Wood,Andrew R.[1], Weyant,Robert J.[8], Segrè, Ayellet V.[11,14,15], Speliotes,Elizabeth K.[10,16], Wheeler,Eleanor[17], Soranzo,Nicole[17,18], Park,Ju-Hyun[7], Yang,Jian[19], Gudbjartsson,Daniel[20], Heard-Costa,Nancy L.[21], Randall,Joshua C.[13], Qi,Lu[22,23], Smith,Albert Vernon[24,25], Mägi,Reedik[13], Pastinen,Tomi[26,27,28], Liang,Liming[29], Heid,Iris M.[30,31], Luan,Jian'an[32], Thorleifsson,Gudmar[20], Winkler,Thomas W.[30], Goddard,Michael E.[33,34], Lo,Ken Sin[5], Palmer,Cameron[9,10], Workalemahu,Tsegaselassie[22], Aulchenko,Yurii S.[2,4], Johansson,Åsa[35,36], Zillikens,M.Carola[3], Feitosa,Mary F.[37], Esko,Tõnu[38,39,40], Johnson,Toby[41,42,43,44], Ketkar,Shamika[37], Kraft,Peter[45,46], Mangino,Massimo[18], Prokopenko,Inga[13,47], Absher,Devin[48] , Albrecht,Eva[31], Ernst,Florian[49], Glazer,Nicole L.[50], Hayward,Caroline[51], Hottenga,Jouke-Jan[52], Jacobs,Kevin B.[53], Knowles,Joshua W.[54], Kutalik,Zoltán[41,42], Monda,Keri L.[55], Polasek,Ozren[56,57], Preuss,Michael[58], Rayner,Nigel W.[13,47], Robertson,Neil R.[13,47], Steinthorsdottir,Valgerdur[20], Tyrer,Jonathan P.[59], Voight,Benjamin F.[11,14,15], Wiklund,Fredrik[60], Xu,Jianfeng[61], Zhao,Jing Hua[32], Nyholt,Dale R.[62], Pellikka,Niina[63,64], Perola,Markus[63,64], Perry,John [R.B].**1**, Surakka,Ida[63,64], Tammesoo,Mari-Liis[38], Altmaier, Elizabeth L.[9,10], Amin,Najaf[2], Aspelund,Thor[24,25], Bhangale,Tushar[65], Boucher,Gabrielle[5], Chasman,Daniel I.[66,67], Chen,Constance[68], Coin,Lachlan[69], Cooper,Matthew N.[70], Dixon,Anna L.[71], Gibson, Quince[72], Grundberg,Elin[17,26,27], Hao,Ke[73], Junttila, M. Juhani[74], Kaplan,Lee M.[16,67,75], Kettunen,Johannes[63,64], König,Inke R.[58], Kwan,Tony[26,27], Lawrence,Robert W.[70], Levinson,Douglas F.[76], Lorentzon,Mattias[77], McKnight,Barbara[78], Morris,Andrew P.[13], Müller,Martina[31,79,80], Ngwa,Julius Suh[81], Purcell,Shaun[14,82,83], Rafelt,Suzanne[84], Salem,Rany M.[9,10], Salvi,Erika[85,86], Sanna,Serena[87],

Shi,Jianxin**7**, Sovio,Ulla[69], Thompson,John R.[88,89],Turchin,Michael C.[9,10], Vandenput,Liesbeth[77], Verlaan,Dominique J.[26,27], Vitart,Veronique[51], White,Charles C.[81], Ziegler,Andreas[90], Almgren,Peter[91], Balmforth,Anthony J.[92], Campbell,Harry[93], Citterio,Lorena[94], De Grandi,Alessandro[95], Dominiczak,Anna[96], Duan,Jubao[97], Elliott,Paul[69], Elosua, Roberto[98], Eriksson,Johan G.[99,100,101,102,103], Freimer,Nelson B.[104], Geus,Eco J.C.[52], Glorioso,Nicola[105], Haiqing,Shen[72], Hartikainen,Anna-Liisa[106], Havulinna,Aki S.[107], Hicks,Andrew A.[95], Hui,Jennie[70,108,109], Igl,Wilmar[35], Illig,Thomas[31], Jula,Antti[110], Kajantie,Eero[100], Kilpeläinen,Tuomas O.[32], Koiranen,Markku[111], Kolcic,Ivana[56], Koskinen,Seppo[107], Kovacs,Peter[112], Laitinen,Jaana[113], Liu,Jianjun[114], Lokki,Marja-Liisa[115], Marusic,Ana[116], Maschio,Andrea[87], Meitinger,Thomas[117,118], Mulas,Antonella[87], Paré,Guillaume[119], Parker,Alex N.[120], Peden,John F.[13,121], Petersmann,Astrid[122], Pichler,Irene[95], Pietiläinen,Kirsi H.[123,124], Pouta,Anneli[106,125], Ridderstråle,Martin[126], Rotter,Jerome I.[127], Sambrook,Jennifer G.[128,129], Sanders,Alan R.[97], Schmidt,Carsten Oliver[130], Sinisalo,Juha[131], Smit,Jan H.[132], Stringham,Heather M.[8], Walters,G.Bragi[20], Widen,Elisabeth[63], Wild,Sarah H.[93], Willemsen,Gonneke[52], Zagato,Laura[94], Zgaga,Lina[56], Zitting,Paavo[133], Alavere,Helene[38], Farrall,Martin[13,121,134], McArdle,Wendy L.[135], Nelis,Mari[38,39,40], Peters,Marjolein J.[3,4], Ripatti,Samuli[63,64], van Meurs,Joyce B.J.[2,3,4], Aben,Katja K.[136], Ardlie,Kristin G[11], Beckmann,Jacques S.[41,137], Beilby,John P.[108,109,138], Bergman,Richard N.[139], Bergmann,Sven[41,42], Collins,Francis S.[140], Cusi,Daniele[85], den Heijer,Martin[141], Eiriksdottir,Gudny[24], Gejman,Pablo V.[97], Hall,Alistair S.[92], Hamsten,Anders[142], Huikuri,Heikki V.[74,74], Iribarren,Carlos[143,144], Kähönen,Mika[145], Kaprio,Jaakko[63,123,146], Kathiresan,Sekar[11,14,147,148,149], Kiemeney,Lambertus[136,150,151], Kocher,Thomas[152], Launer,Lenore J.[153], Lehtimäki,Terho[154], Melander,Olle[126], Mosley Jr,Tom H.[155], Musk,Arthur W.[109,156], Nieminen,Markku S.[131], O'Donnell,Christopher J.[148,157], Ohlsson,Claes[77], Oostra,Ben[158], Palmer,Lyle J.[70,109], Raitakari,Olli[159], Ridker,Paul M.[66,67], Rioux,John D.[5,6], Rissanen,Aila[124, 2] Rivolta,Carlo[41], Schunkert,Heribert[160], Shuldiner,Alan R.[72,161], Siscovick,David S.[162,163], Stumvoll,Michael[164,165], Tönjes,Anke[164,166], Tuomilehto,Jaakko[167,168,169], van Ommen,Gert-Jan[170], Viikari,Jorma[171], Heath,Andrew C.[172], Martin,Nicholas G.[173], Montgomery,Grant W.[174] , Province,Michael A.[37,175], Kayser,Manfred[176], Arnold,Alice M.[78,177], Atwood,Larry D.[21], Boerwinkle,Eric[178], Chanock,Stephen J.[7], Deloukas,Panos[17], Gieger,Christian[31], Grönberg,Henrik[60], Hall,Per[60], Hattersley,Andrew T.[1], Hengstenberg,Christian[179,180], Hoffman,Wolfgang[130], Lathrop,G.Mark[181], Salomaa,Veikko[107], Schreiber,Stefan[182], Uda,Manuela[87], Waterworth,Dawn[183], Wright,Alan F.[51], Assimes,Themistocles L.[54], Barroso,Inês[17,184], Hofman,Albert[2,4], Mohlke,Karen L.[185], Boomsma,Dorret I.[52], Caulfield,Mark J.[44], Cupples,L.Adrienne[81], Erdmann,Jeanette[160], Fox,Caroline S.[186], Gudnason,Vilmundur[24,25], Gyllensten,Ulf[35], Harris,Tamara B.[153], Hayes,Richard B.[187], Jarvelin,Marjo-Riitta[69,111,125,188], Mooser,Vincent[183], Munroe,Patricia B.[44], Ouwehand,Willem H.[17,128,129], Penninx,Brenda W.[132,189,190], Pramstaller,Peter P.[95,191,192], Quertermous,Thomas[54], Rudan,Igor[51,116] , Samani,Nilesh J.[84,88], Spector,Timothy D.[18], Völzke,Henry[130], Watkins, Hugh on behalf of Procardis Consortium[13,121], Wilson,James F.[93], Groop,Leif C.[91], Haritunians,Talin[127], Hu,Frank B.[22,23,45], Kaplan,Robert C.[193], Metspalu,Andres[38,39,40], North,Kari E.[55,194], Schlessinger,David[195], Wareham,Nicholas J.[32], Hunter,David J.[22,23,45], O'Connell,Jeffrey R.[72], Strachan,David P.**1**[96],Wichmann,H.-

Erich[31,80,197], Borecki,Ingrid B.[37,175], van Duijn,Cornelia M.[2,4], Schadt,Eric E.[198,199], Thorsteinsdottir,Unnur[20,200], Peltonen,Leena[17,63,64,82,201], Uitterlinden,André[2,3,4], Visscher,Peter M.[19], Chatterjee,Nilanjan[7], Loos,Ruth J.F.[32], Boehnke,Michael[8], McCarthy,Mark I.[13,47,202], Ingelsson,Erik[60], Lindgren,Cecilia M.[13,47], Abecasis,Gonçalo R.[8], Stefansson,Kari[20,200], Frayling,Timothy M.[1], Hirschhorn,Joel N.[9,10,203]

1. Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, EX1 2LU, UK
2. Department of Epidemiology, Erasmus MC, Rotterdam, 3015GE, The Netherlands
3. Department of Internal Medicine, Erasmus MC, Rotterdam, 3015GE, The Netherlands
4. Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA)
5. Montreal Heart Institute, Montreal, Quebec, H1T 1C8, Canada
6. Department of Medicine, Université de Montréal, Montreal, Quebec, H3T 1J4, Canada
7. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892, USA
8. Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA
9. Divisions of Genetics and Endocrinology and Program in Genomics, Children's Hospital, Boston, Massachusetts 02115, USA
10. Metabolism Initiative and Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA
11. Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA
12. Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115 USA
13. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK
14. Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.
15. Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA
16. Division of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA
17. Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK
18. Department of Twin Research and Genetic Epidemiology, King's College London, Lambeth Palace Rd, London, SE1 7EH, UK
19. Queensland Statistical Genetics Laboratory, Queensland Institute of Medical Research, Queensland 4006, Australia
20. deCODE Genetics, 101 Reykjavik, Iceland
21. Department of Neurology, Boston University School of Medicine, Boston, Massachusetts 02118, USA
22. Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts 02115, USA
23. Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA
24. Icelandic Heart Association, Kopavogur, Iceland
25. University of Iceland, Reykjavik, Iceland
26. McGill University and Genome Québec Innovation Centre, Montréal, Québec H3A 1A4, Canada.
27. Department of Human Genetics, McGill University Health Centre, McGill University, Montréal, Québec H3G 1A4, Canada
28. Department of Medical Genetics, McGill University Health Centre, McGill University, Montréal, Québec H3G 1A4, Canada
29. Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Cambridge, Massachusetts 02138, USA
30. Regensburg University Medical Center, Department of Epidemiology and Preventive Medicine, 93053 Regensburg, Germany
31. Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany
32. MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK
33. University of Melbourne, Parkville 3010, Australia
34. Department of Primary Industries, Melbourne, Victoria 3001, Australia
35. Department of Genetics and Pathology, Rudbeck Laboratory, University of Uppsala, SE-75185 Uppsala, Sweden
36. Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, N-7489, Norway
37. Department of Genetics, Washington University School of Medicine, St Louis, Missouri 63110, USA
38. Estonian Genome Center, University of Tartu, Tartu 50410, Estonia
39. Estonian Biocenter, Tartu 51010, Estonia
40. Institute of Molecular and Cell Biology, University of Tartu, Tartu 51010, Estonia
41. Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland
42. Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland
43. Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary, University of London, London, UK
44. Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK
45. Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA
46. Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, USA
47. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, OX3 7LJ, UK
48. Hudson Alpha Institute for Biotechnology, Huntsville, Alabama 35806, USA

49. Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt-University Greifswald, 17487 Greifswald, Germany
50. Cardiovascular Health Resarch Unit and Department of Medicine, University of Washington, Seattle, Washington 98101, USA
51. MRC Human Genetics Unit, Institute for Genetics and Molecular Medicine, Western General Hospital, Edinburgh, EH4 2XU, Scotland, UK
52. Department of Biological Psychology, VU University Amsterdam, 1081 BT Amsterdam, The Netherlands
53. Core Genotyping Facility, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702, USA
54. Department of Medicine, Stanford University School of Medicine, Stanford, California 94305, USA
55. Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514 USA
56. Andrija Stampar School of Public Health, Medical School, University of Zagreb, 10000 Zagreb, Croatia
57. Gen-Info Ltd, 10000 Zagreb, Croatia
58. Universität zu Lübeck, Institut für Medizinische Biometrie und Statistik, 23562 Lübeck, Germany
59. Department of Oncology, University of Cambridge, Cambridge, CB1 8RN, UK
60. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden
61. Center for Human Genomics, Wake Forest University, Winston-Salem, North Carolina 27157, USA
62. Neurogenetics Laboratory, Queensland Institute of Medical Research, Queensland 4006, Australia
63. Institute for Molecular Medicine Finland (FIMM), University of Helsinki, 00014, Helsinki, Finland
64. National Institute for Health and Welfare, Department of Chronic Disease Prevention, Unit of Public Health Genomics, 00014, Helsinki, Finland
65. Department of Genome Sciences, University of Washington, Seattle, 98195 Washington, USA
66. Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts 02215, USA
67. Harvard Medical School, Boston, Massachusetts 02115, USA
68. Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA
69. Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London,W2 1PG , UK
70. Centre for Genetic Epidemiology and Biostatistics, University of Western Australia, Crawley, Western Australia 6009, Australia
71. Royal National Hospital for Rheumatic Diseases and University of Bath, Bath, BA1 1RL, UK
72. Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA
73. Genetics Department, Rosetta Inpharmatics, a Wholly Owned Subsidiary of Merck & Co. Inc., Seattle, Washington 98109, USA
74. Department of Internal Medicine, University of Oulu, 90014 Oulu, Finland
75. MGH Weight Center, Massachusetts General Hospital, Boston, Massachusetts 02114, USA
76. Stanford University School of Medicine, Stanford, California 93405, USA
77. Department of Internal Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, 413 45 Gothenburg, Sweden
78. Departments of Biostatistics, University of Washington, Seattle, Washington 98195, USA
79. Ludwig-Maximilians-University, Department of Medicine I, University Hospital Grosshadern, 81377 Munich, Germany
80. Ludwig-Maximilians-Universität, Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, 81377 Munich, Germany
81. Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts 02118, USA
82. The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA
83. Department of Psychiatry, Harvard Medical School, Boston, Massachusetts 02115, USA
84. Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, LE3 9QP, UK
85. University of Milan, Department of Medicine, Surgery and Dentistry, 20139 Milano, Italy
86. KOS Genetic Srl, 20123 Milan, Italy
87. Istituto di Neurogenetica e Neurofarmacologia del CNR, Monserrato, 09042, Cagliari, Italy
88. Leicester NIHR Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, LE3 9QP, UK
89. Department of Health Sciences, University of Leicester, University Road, Leicester, LE1 7RH, UK
90. Universität zu Lübeck, Institut für Medizinische Biometrie und Statistik, 23562 Lübeck, Germany
91. Lund University Diabetes Centre, Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden
92. Multidisciplinary Cardiovascular Research Centre (MCRC), Leeds Institute of Genetics, Health and Therapeutics (LIGHT), University of Leeds, Leeds LS2 9JT, UK
93. Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland
94. University Vita-Salute San Raffaele, Division of Nephrology and Dialysis, 20132 Milan, Italy
95. Institute of Genetic Medicine, European Academy Bozen/Bolzano (EURAC), Bolzano/Bozen, 39100, Italy. Affiliated Institute of the University of Lübeck, Lübeck, Germany.
96. British Heart Foundation Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, G12 8TA, UK
97. Northshore University Healthsystem, Evanston, Ilinois 60201, USA
98. Cardiovascular Epidemiology and Genetics, Institut Municipal D'investigacio Medica and CIBER Epidemiología y Salud Pública, Barcelona, Spain
99. Department of General Practice and Primary health Care, University of Helsinki, Helsinki, Finland
100. National Institute for Health and Welfare, 00271 Helsinki, Finland
101. Helsinki University Central Hospital, Unit of General Practice, 00280 Helsinki, Finland
102. Folkhalsan Research Centre, 00250 Helsinki, Finland
103. Vasa Central Hospital, 65130 Vasa, Finland
104. Center for Neurobehavioral Genetics, University of California, Los Angeles, California 90095, USA
105. Hypertension and Cardiovascular Prevention Center, University of Sassari, 07100 Sassari, Italy
106. Department of Clinical Sciences/Obstetrics and Gynecology, University of Oulu, 90014 Oulu, Finland

107. National Institute for Health and Welfare, Department of Chronic Disease Prevention, Chronic Disease Epidemiology and Prevention Unit, 00014, Helsinki, Finland
108. PathWest Laboratory of Western Australia, Department of Molecular Genetics, J Block, QEII Medical Centre, Nedlands, Western Australia 6009, Australia
109. Busselton Population Medical Research Foundation Inc., Sir Charles Gairdner Hospital, Nedlands, Western Australia 6009, Australia
110. National Institute for Health and Welfare, Department of Chronic Disease Prevention, Population Studies Unit, 20720 Turku, Finland
111. Institute of Health Sciences, University of Oulu, 90014 Oulu, Finland
112. Interdisciplinary Centre for Clinical Research, University of Leipzig, 04103 Leipzig, Germany
113. Finnish Institute of Occupational Health, 90220 Oulu, Finland
114. Human Genetics, Genome Institute of Singapore, Singapore 138672, Singapore
115. Transplantation Laboratory, Haartman Institute, University of Helsinki, 00014, Helsinki, Finland
116. Croatian Centre for Global Health, School of Medicine, University of Split, Split 21000, Croatia
117. Institute of Human Genetics, Klinikum rechts der Isar der Technischen Universität München, 81675 Munich, Germany
118. Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany
119. Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario L8N3Z5, Canada
120. Amgen, Cambridge, Massachusetts 02139, USA
121. Department of Cardiovascular Medicine, University of Oxford, Level 6 West Wing, John Radcliffe Hospital, Headley Way, Headington, Oxford, OX3 9DU
122. Institut für Klinische Chemie und Laboratoriumsmedizin, Universität Greifswald, 17475 Greifswald, Germany
123. Finnish Twin Cohort Study, Department of Public Health, University of Helsinki, 00014, Helsinki, Finland
124. Obesity Research unit, Department of Psychiatry, Helsinki University Central Hospital, Helsinki, Finland
125. National Institute for Health and Welfare, 90101 Oulu, Finland
126. Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden
127. Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA
128. Department of Haematology, University of Cambridge, Cambridge CB2 0PT, UK
129. NHS Blood and Transplant, Cambridge Centre, Cambridge, CB2 0PT, UK
130. Institut für Community Medicine, 17489 Greifswald, Germany
131. Division of Cardiology, Cardiovascular Laboratory, Helsinki University Central Hospital, 00029 Helsinki, Finland
132. Department of Psychiatry/EMGO Institute, VU University Medical Center, 1081 BT Amsterdam, The Netherlands
133. Department of Psychiatrics, Lapland Central Hospital, 96101 Rovaniemi, Finland
134. Cardiovascular Medicine, University of Oxford, Wellcome Trust Centre for Human Genetics, Oxford, OX3 7BN, UK
135. Avon Longitudinal Study of Parents and Children (ALSPAC) Laboratory, Department of Social Medicine, University of Bristol, Bristol, BS8 2BN, UK
136. Comprehensive Cancer Center East, 6501 BG Nijmegen, The Netherlands
137. Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois (CHUV) University Hospital, 1011 Lausanne, Switzerland
138. School of Pathology and Laboratory Medicine, University of Western Australia, Nedlands, Western Australia 6009,Australia
139. Department of Physiology and Biophysics, Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA
140. National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA
141. Department of Endocrinology, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands
142. Atherosclerosis Research Unit, Department of Medicine, Solna,Karolinska Institutet, Karolinska University Hospital, 171 76 Stockholm, Sweden
143. Division of Research, Kaiser Permanente Northern California, Oakland, California 94612, USA
144. Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California 94107, USA
145. Department of Clinical Physiology, University of Tampere and Tampere University Hospital, 33520 Tampere, Finland;
146. National Institute for Health and Welfare, Department of Mental Health and Substance Abuse Services, Unit for Child and Adolescent Mental Health, 00271 Helsinki, Finland
147. Cardiovascular Research Center and Cardiology Division, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.
148. Framingham Heart Study of the National, Heart, Lung, and Blood Institute and Boston University, Framingham, Massachusetts 01702, USA
149. Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA
150. Department of Epidemiology, Biostatistics and HTA, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands
151. Department of Urology, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands
152. Zentrum für Zahn-, Mund- und Kieferheilkunde, 17489 Greifswald, Germany
153. Laboratory of Epidemiology, Demography, Biometry, National Institute on Aging, National Institutes of Health, Bethesda, Maryland 20892, USA
154. Department of Clinical Chemistry, University of Tampere and Tampere University Hospital, 33520 Tampere, Finland
155. Department of Medicine, Division of Geriatrics, University of Mississippi Medical Center, Jackson, Mississippi 39216, USA
156. School of Medicine and Pharmacology, University of Western Australia, Perth, Western Australia 6009, Australia
157. National, Lung, and Blood Institute, National Institutes of Health, Framingham, Massachusetts 01702, USA
158. Department of Clinical Genetics, Erasmus MC, Rotterdam, 3015GE, The Netherlands
159. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku and the Department of Clinical Physiology, Turku University Hospital, 20520 Turku, Finland
160. Universität zu Lübeck, Medizinische Klinik II, 23562 Lübeck, Germany

161. Geriatrics Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, Maryland 21201, USA
162. Cardiovascular Health Research Unit, University of Washington, Seattle, Washington 98101, USA
163. Departments of Medicine and Epidemiology, University of Washington, Seattle, Washington 98195, USA
164. Department of Medicine, University of Leipzig, 04103 Leipzig, Germany
165. LIFE Study Centre, University of Leipzig, Leipzig, Germany
166. Coordination Centre for Clinical Trials, University of Leipzig, Härtelstr. 16-18, 04103 Leipzig, Germany
167. National Institute for Health and Welfare, Diabetes Prevention Unit, 00271 Helsinki, Finland
168. Hjelt Institute, Department of Public Health, University of Helsinki, 00014 Helsinki, Finland
169. South Ostrobothnia Central Hospital, 60220 Seinajoki, Finland
170. Department of Human Genetics and Center of Medical Systems Biology, Leiden University Medical Center, 2333 ZC Leiden, the Netherlands
171. Department of Medicine, University of Turku and Turku University Hospital, 20520 Turku, Finland
172. Department of Psychiatry and Midwest Alcoholism Research Center, Washington University School of Medicine, St Louis, Missouri 63108, USA
173. Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Queensland 4006, Australia
174. Molecular Epidemiology Laboratory, Queensland Institute of Medical Research, Queensland 4006, Australia
175. Division of Biostatistics,Washington University School of Medicine,St.Louis, Missouri 63110, USA
176. Department of Forensic Molecular Biology, Erasmus MC, Rotterdam, 3015GE, The Netherlands
177. Collaborative Health Studies Coordinating Center, Seattle, Washington 98115, USA
178. Human Genetics Center and Institute of Molecular Medicine and Division of Epidemiology, University of Texas Health Science Center, Houston, Texas 77030, USA
179. Klinik und Poliklinik für Innere Medizin II, Universität Regensburg, 93053 Regensburg, Germany
180. Regensburg University Medical Center, Innere Medizin II, 93053 Regensburg, Germany
181. Centre National de Genotypage, Evry, Paris 91057, France
182. Christian-Albrechts-University, University Hospital Schleswig-Holstein, Institute for Clinical Molecular Biology and Department of Internal Medicine I, Schittenhelmstrasse 12, 24105 Kiel
183. Genetics Division, GlaxoSmithKline, King of Prussia, Pennsylvania 19406, USA
184. University of Cambridge Metabolic Research Labs, Institute of Metabolic Science Addenbrooke's Hospital, CB2 OQQ, Cambridge, UK
185. Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599, USA
186. Division of Intramural Research, National Heart, Lung and Blood Institute, Framingham Heart Study, Framingham, Massachusetts 01702, USA
187. New York University Medical Center, New York, New York 10016, USA
188. Biocenter Oulu, University of Oulu, 90014 Oulu, Finland
189. Department of Psychiatry, Leiden University Medical Centre, 2300 RC Leiden, The Netherlands
190. Department of Psychiatry, University Medical Centre Groningen, 9713 GZ Groningen, The Netherlands
191. Department of Neurology, General Central Hospital, Bolzano, Italy
192. Department of Neurology, University of Lübeck, Lübeck, Germany.
193. Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York 10461, USA
194. Carolina Center for Genome Sciences, School of Public Health, University of North Carolina Chapel Hill, Chapel Hill, North Carolina 27514, USA
195. Laboratory of Genetics, National Institute on Aging, Baltimore, Maryland 21224, USA
196. Division of Community Health Sciences, St George's, University of London, London, SW17 0RE, UK
197. Klinikum Grosshadern, 81377 Munich, Germany
198. Pacific Biosciences, Menlo Park, California 94025, USA
199. Sage Bionetworks, Seattle, Washington 98109, USA
200. Faculty of Medicine, University of Iceland, 101 Reykjavík, Iceland
201. Department of Medical Genetics, University of Helsinki, 00014 Helsinki, Finland
202. NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford, OX3 7LJ, UK
203. Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

FINANCIAL DISCLOSURES

Patrick Sulem, Daniel F. Gudbjartsson, Kari Stefansson and Unnur Thorsteinsdottir are employed by deCODE Genetics; Valur Emilsson is employed by Merck & Co; Vincent Mooser and Dawn M. Waterworth are employed by GlaxoSmithKline; Alex Parker and Kim Tsui are employed by Amgen. Inês Barroso and spouse own stock in Incyte Ltd and GlaxoSmithKline.

AUTHOR CONTRIBUTIONS

Writing group: DIB, HAB, DIC, DLC, LC, EWD, MJE, CEE (lead), TE, BF, NF, FG, DFG, CH, TBH, DJH, DLK, KKL, MMan, MMar, JMM (chair), Ame, AMu, KKO (lead), JRBP, LS, EAS, PS, UT, AGU, CMvD, SWvW, JAV, EW, GZ

Analysis working group: CEE, AM, KKO, JRBP, PS

Secondary analyses: Jenny A.Visser, John R.B. Perry, Andrew D. Johnson, Daniel Levy, Andrew S. Plump, Valur Emilsson, Ayellet V. Segrè

Contributing cohorts:

	Oversight	Analytical lead	Phenotype and genotyping
AGES	Tamara B. Harris	Albert V. Smith	Thor Aspelund, Gudny Eiriksdottir, Vilmundur Gudnason, Lenore J Launer, Melissa Garcia, Michael A. Nalls
Amish	Elizabeth A. Streeten	Patrick F. McArdle	Alan R. Shuldiner
ARIC	Ellen W. Demerath	Nora Franceschini	Eric Boerwinkle, David Couper, Aaron R. Folsom
B58C	David P Strachan	Cathy E Elks	Wendy L. McArdle
CoLaus	Vincent Mooser, Dawn M Waterworth	Zoltán Kutalik	Gerard Waeber, Peter Vollenweider, Sven Bergmann
deCODE	Kari Stefansson, Unnur Thorsteinsdottir	Patrick Sulem, Daniel F Gudbjartsson	Thorunn Rafnar, Laufey Tryggvadottir
DNBC	Heather A Boyd	Bjarke Feenstra	Frank Geller, Mads Melbye, Jeffrey C Murray
EGCUT	Andres Metspalu	Tõnu Esko	Helene Alavere, Mari Nelis, Mari-Liis Tammesoo
EPIC-Norfolk	Kay-Tee Khaw, Nicholas J Wareham	Ruth JF Loos, Jing Hua Zhao	Ines Barroso, Panos Deloukas, Tuomas O. Kilpeläinen, Shengxu Li, Ken K Ong
ERF	Cornelia M. van Duijn	Sophie H. van Wingerden	Najaf Amin, Ben A. Oostra
FHS	Joanne M. Murabito	Kathyrn L. Lunetta	Andrea D. Coviello, David Karasik, Douglas P Kiel, Joanne M. Murabito, Wei Vivian Zhuang
HBCS	Elisabeth Widen	Diana L Cousminer	Johan Eriksson, Aarno Palotie, Leena Peltonen, Emmi Tikkanen

Health 2000 (Genmets)	Elisabeth Widen	Diana L Cousminer	Leena Peltonen, Veikko Salomaa, Emmi Tikkanen
InCHIANTI	Anna Murray	John RB Perry	Stefania Bandinelli, Luigi Ferrucci, Dena G. Hernandez, Michael N Weedon
Indiana	Michael J. Econs	Daniel L. Koller	Tatiana Foroud, Munro Peacock
NFBC	Elisabeth Widen	Diana L Cousminer	Marjo-Riitta Järvelin, Leena Peltonen, Anneli Pouta, Ulla Sovio
NHS	David J. Hunter	Chunyan He	Stephen J. Chanock, Marilyn C. Cornelis, Susan E. Hankinson, Frank B. Hu, Peter Kraft, Rob M. Van Dam
NTR	Dorret I Boomsma	Jouke- Jan Hottenga	Eco JC de Geus, Gonneke Willemsen
QIMR	Grant W Montgomery	Enda M Byrne	Andrew C Heath, Nicholas G Martin
RS	Albert Hofman, André G Uitterlinden	Lisette Stolk, Jenny A Visser	Joop SE Laven, Fernando Rivadeneira, Joyce BJ van Meurs
SAGE	Laura J. Bierut	Peng Lin	John P. Rice
SardiNIA	Mara Marongiu, Laura Crisponi	Serena Sanna, Eleonora Porcu	Fabio Busonero, Liana Ferreli, David Schlessinger, Angelo Scuteri, Manuela Uda
TwinsUK	Tim D Spector	Massimo Mangino	Hannah Blackburn, Nicole Soranzo, Guangju Zhai
WGHS	Paul M Ridker	Daniel I Chasman	Julie E Buring, Guillaume Paré, Alex N Parker, Kim Tsui, Lauren Young
ALSPAC	George Davey Smith	Kate Northstone	Andrew R Ness, Susan M. Ring
BHS	Sarah S. Murray, Nicholas J. Schork	Erin N. Smith	Gerald S. Berenson, Wei Chen, Sathanur R. Srinivasan,
INGI	Paolo Gasparini, Daniela Toniolo	Tanguy Corre, Sheila Ulivi	Pio d'Adamo, Cinzia Sala
KORA	H.-Erich Wichmann	Eva Albrecht	Angela Döring, Christian Gieger, Thomas Illig
SASBAC	Per Hall	Erik Ingelsson	Jianjun Liu
SEARCH	Paul Pharoah, Douglas F Easton	Cathy E Elks	Jonathon Tyrer
STR_MZ twins	Ulf de Faire	Patrik KE Magnusson	Nancy L. Pedersen
Raine	Craig E. Pennell	Nicole M. Warrington	Lyle J. Palmer, Craig E. Pennell
Orcades	James F Wilson	Lina Zgaga	Harry Campbell
SPLIT	Igor Rudan, Alan F. Wright	Caroline Hayward	Ozren Polasek, Ivana Kolcic, Pau Navarro
KORCULA	Igor Rudan, Alan F. Wright	Caroline Hayward	Ozren Polasek, Ivana Kolcic, Pau Navarro
VIS	Igor Rudan, Alan F. Wright	Caroline Hayward	Ozren Polasek, Ivana Kolcic, Pau Navarro

List of Authors' Institutions

[1] Medical Research Council (MRC) Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK
[2] Genetics of Complex Traits, Peninsula Medical School, University of Exeter, UK
[3] deCODE Genetics, Reykjavik, Iceland
[4] Division of Preventive Medicine, Brigham and Women's Hospital, 900 Commonwealth Avenue East, Boston MA 02215, USA
[5] Harvard Medical School, Boston, Massachusetts, USA
[6] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[7] Department of Public Health, Indiana University School of Medicine, Indiana, USA
[8] Melvin and Bren Simon Cancer Center, Indiana University, Indiana, USA
[9] The National Heart Lung and Blood Institute's Framingham Heart Study, Framingham, MA, USA
[10] Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA
[11] Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands
[12] Queensland Statistical Genetics, Queensland Institute of Medical Research, Brisbane, Australia
[13] The University of Queensland, Brisbane, Australia
[14] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland
[15] Estonian Genome Center, University of Tartu, Tartu, Estonia
[16] Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia
[17] Genotyping Core Facility, Estonian Biocenter, Tartu, Estonia
[18] Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark
[19] Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands
[20] Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indiana, USA
[21] Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland
[22] Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland
[23] Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA
[24] Department of Twin Research and Genetic Epidemiology, King's College London, London, UK
[25] Istituto di Neurogenetica e Neurofarmacologia, Consiglio Nazionale delle Ricerche, Cagliari, Italy
[26] Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, USA
[27] Icelandic Heart Association, Kopavogur, Iceland
[28] University of Iceland, Reykjavik, Iceland
[29] Netherlands Consortium of Healthy Aging, Rotterdam, the Netherlands
[30] Genetic-Epidemiology Unit, Department of Epidemiology and Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands
[31] Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany
[32] Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Milan, Italy
[33] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[34] MRC Human Genetics Unit; Institute of Genetics and Molecular Medicine, Western General Hospital; Edinburgh, UK
[35] Scripps Genomic Medicine, The Scripps Translational Science Institute, and The Scripps Research Institute, La Jolla, CA, USA
[36] Medical Genetics, Department of Reproductive Sciences and Development, University of Trieste, Trieste, Italy
[37] Centre for Genetic Epidemiology and Biostatistics University of Western Australia, Australia
[38] Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland
[39] Geriatric Unit, Azienda Sanitaria di Firenze, Florence, Italy
[40] Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK
[41] Tulane University, New Orleans, LA, USA
[42] Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, USA

[43] Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA
[44] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD 20892, USA
[45] Department of Nutrition, Harvard School of Public Health, Boston, MA, USA
[46] Collaborative Studies Coordinating Center, Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[47] Sections of General Internal Medicine, Preventive Medicine and Endocrinology, Department of Medicine, Boston University School of Medicine, Boston, MA, USA.
[48] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[49] MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, BS8 2BN, UK
[50] Centre for Cancer Genetic Epidemiology, Departments of Oncology and Public Health and Primary Care, University of Cambridge, Cambridge, UK
[51] MPRI, Merck & Co., Inc, 126 Lincoln Ave, Rahway, NJ 07065, USA
[52] National Institute for Health and Welfare, Finland
[53] Department of General Practice and Primary health Care, University of Helsinki, Finland
[54] Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland
[55] Folkhalsan Research Centre, Helsinki, Finland
[56] Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging, Baltimore, Maryland, USA
[57] Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA
[58] Laboratory of Epidemiology, Demography, and Biometry, Intramural Research Program, National Institute on Aging, Bethesda, Maryland, USA
[59] A full list of members is provided in the **Supplementary Note**
[60] Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston, Massachusetts, USA
[61] Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA
[62] Laboratory of Neurogenetics, National Institute of Ageing, Bethesda, MD, USA
[63] Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands
[64] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK
[65] NHLBI Center for Population Studies, Bethesda, MD, USA
[66] Hebrew SeniorLife Institute for Aging Research and Harvard Medical School, Boston, MA, USA
[67] Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge, CB2 0QQ, UK
[68] Medical School; University of Zagreb; Zagreb, 10000; Croatia
[69] Department of Obstetrics and Gynaecology, Erasmus MC, Rotterdam, the Netherlands
[70] Human Genetics, Genome Institute of Singapore, Singapore
[71] Division of Cardiology, Boston University School of Medicine, USA
[72] Genetic Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia
[73] Avon Longitudinal Study of Parents and Children (ALSPAC), Department of Social Medicine, University of Bristol, BS8 2BN, UK
[74] Genetics Division, GlaxoSmithKline, King of Prussia, Pennsylvania, USA
[75] Department of Pediatrics, University of Iowa, Iowa City, IA, USA
[76] Laboratory of Neurogenetics, Intramural Research Program, National Institute on Aging, Bethesda, Maryland, USA
[77] Department of Oral and Dental Science, University of Bristol, BS1 2LY, UK
[78] Department of Medicine, Indiana University School of Medicine, Indiana, USA
[79] Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA
[80] Genetic and Molecular Epidemiology Laboratory, McMaster University, 1200 Main St. W MDCL Rm. 3206, Hamilton, ON, L8N3Z5, Canada
[81] Amgen, 1 Kendall Square, Building 100, Cambridge, MA 02139, USA

[82] Deceased
[83] School of Women's and Infants' Health, The University of Western Australia, Australia
[84] Gen Info Ltd; Zagreb, 10000; Croatia
[85] Cardiovascular Disease, Merck Research Laboratory, Rahway, NJ 07065, USA
[86] Croatian Centre for Global Health; University of Split Medical School; Split, 21000; Croatia
[87] Gerontology Research Center, National Institute on Aging, Baltimore, Maryland, USA
[88] UOC Geriatria - Istituto Nazionale Ricovero e Cura per Anziani IRCCS – Rome, Italy
[89] Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, USA
[90] Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, Baltimore, Maryland, USA
[91] Division of Community Health Sciences, St. George's, University of London, London, UK
[92] Icelandic Cancer Registry, Reykjavik, Iceland
[93] Departments of Epidemiology and Public Health and Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[94] Department of Internal Medicine, BH-10 Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland
[95] Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany
[96] Klinikum Grosshadern, Munich, Germany
[97] Molecular Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia
[98] Division of Cardiology, Brigham and Women's Hospital
[99] Faculty of Medicine, University of Iceland, Reykjavik, Iceland
[100] Department of Paediatrics, University of Cambridge, Cambridge, UK